



EP 0 786 519 A2

|    |             |            |            |             |             |             |      |
|----|-------------|------------|------------|-------------|-------------|-------------|------|
|    | GTCTGTAATT  | ACAAAATCAT | CTTCTTGACG | AAGTAAAGAT  | TCTAACAAAA  | CACGAATTGA  | 3060 |
|    | ATAAGGTAAA  | TTGGAAACTT | TAGTAATACC | TTGCTCTTCT  | ACAGCTTTTA  | AATCATAGTA  | 3120 |
| 5  | AGTATAACTT  | TGGCCATTCA | AGTCAAAATG | TTTTTTTGAT  | TGCTCTTTAA  | AATTGTCAGC  | 3180 |
|    | CATTTAATGA  | TCCCCCTTGA | TACATTTTTA | TATTTATATG  | CCTTGATTAA  | ATTGTATTAT  | 3240 |
|    | TATATTTAT   | GATAACAAC  | TCATCATGCT | TAGAAAAAGC  | TTAATTTAGG  | TTTTGACTTT  | 3300 |
| 10 | TTAATCAGAG  | TATATAAGCA | AAACTTATCA | TACAGGTAAG  | GTGTAATAAG  | TATTTTTTAT  | 3360 |
|    | TAATTGAGAA  | TAATTATCAA | TTTCGCGAAT | GATTCAATTC  | AATTTTTAAA  | CGTATTATTT  | 3420 |
|    | CATTGAGCAG  | AAAGAAAATT | ATGGCACCAA | ACTTTAATAT  | TTTTTTCAAT  | GTCAATCTTT  | 3480 |
| 15 | TGATGGGAGT  | GGGACAGAAA | TGATATTTTC | GCAAAATTTA  | TTTCGTCGTC  | CCACCCCAAC  | 3540 |
|    | TTGCATTGTC  | TGTAGAAATT | GGGAATCCAA | TTTCTCTTTG  | TTGGGGCCCA  | TCCCAACTT   | 3600 |
|    | GCACATTATT  | GTAAGCTGAC | TTTTCGTGAG | CTTCTGTGTT  | GGGGCCCTCA  | CCCCAACTCG  | 3660 |
| 20 | CATTGCCTGT  | AGAATTCTT  | TTCGAAATTC | TCTGTGTGG   | GGCCCCGAC   | TAGAATTGAA  | 3720 |
|    | AAAAGCTTGT  | TACAAGCGCA | TTTTCGTTCA | GTCACTACT   | GCCAATATAA  | CTTCGTAGAG  | 3780 |
| 25 | CATAGAAAT   | TGATTTATGT | CCGACCTGA  | GTTAATTTTC  | TATAAAAGTA  | TATTTAAITTT | 3840 |
|    | GCGTTTATAC  | CGTCAAACTT | CACCTTAGCT | TTGTCAAACC  | CCTTCTTATT  | AAGTTTTCAG  | 3900 |
|    | AAATAAACCT  | ATCTTAAAT  | ATAAAAAAAT | CGAGAATTCG  | TAGTTTAAATA | ACGAAATTTCT | 3960 |
| 30 | CGTCTTATC   | CTTTTGAATA | TACTCAATTT | TCCACAAAAA  | CAAAACAAGTA | GTATATCTGT  | 4020 |
|    | TCTAGCTACT  | AGAATGACAT | ACTACTTGTT | ATTAATAATC  | TTAACTAAAC  | TTTATTAGTT  | 4080 |
|    | ATCTTTTTTC  | TCTATATTTT | TACGTGACTG | ACGCTTTTCA  | AGAATGTGCA  | ATTGTAATTC  | 4140 |
| 35 | TTCTTGTTGA  | CTCTTGATAT | ATTCTTGTA  | GCGATGTTTA  | TTGGGAGTCA  | ATGTTAAACC  | 4200 |
|    | TAGGAATTTA  | CGTTCCTGGT | TGCGATCCTT | GTAGAAACTT  | ACCATCATGA  | GTATGACGAC  | 4260 |
| 40 | AAAGGAGAA   | GGGAATGCAC | TTATAATTGC | AGCACTTTGA  | ATCGCAATTIA | AAGCTTCAGC  | 4320 |
|    | GCGGTTACCG  | CCACCAGCTA | ATAAAGTAC  | AAATGCTATT  | AAGGCTCTGT  | AAATCCCCCA  | 4380 |
|    | AACAACCTTT  | ACCATACTAG | ATGGATTTTA | TGAACCAAT   | GTGTTTGTCA  | TTCTTAATAC  | 4440 |
| 45 | AAATGTTGCT  | GAGTCAGCAG | ATGTAATAAA | GAATGATGCA  | ATTAATAATA  | ATGCAATCAA  | 4500 |
|    | CGATAAAACA  | ATGCCAAATG | GCACATGATT | AAACACTCCA  | AATAGCTGTG  | TTTCAGGAGT  | 4560 |
|    | CATATCAAAA  | ATTTCTTTGT | GTTCCTTACC | TGTCGTGAT   | CCTAATACAC  | CAAAGACACT  | 4620 |
| 50 | AAACCAAAACA | AAACTAACAA | TTGCTGGAAC | TAGCAAGACA  | CCAGAAATGA  | ACTCTCTAAT  | 4680 |
|    | TGAACGTCTCT | TTTGAAACTC | GTGCAATAAA | CACCTCCAACG | AATGGACTCC  | AACTTAACCA  | 4740 |

55

|    |  |      |
|----|--|------|
|    | TGCTGTATCA AAACATTAA ACAAGATGT GTTTAGTAAA CTACCCGTAG AGCTAGTTAA    | 4860 |
|    | CATATTTAAA ATAAGAACAG TTGGTCCAAC AATTAAAGCA GCTACCATTAA AAATAGTACC | 4920 |
| 5  | TAAACCAATG TTCAAGTTAC TTAAGTATTG AATACCTTTA CTTAATCCAG ACCATGCACT  | 4980 |
|    | TGCTATAAAT AAGTAGTAA CAACAATGAT GATAATCGCT TGTACAAAGC TATTGTTTGG   | 5040 |
|    | AACATTGAAC AAGTAATGTA AACCACCATT AATTTGTAGA GCACCCATAC CTAACGAAAC  | 5100 |
| 10 | GGCTACCCCA ACGATTGTCG CAAATACAGA TAAACGTCA ATAAAAATCC CAATAGGACC   | 5160 |
|    | TTCTACTTTA TCACCTAAAA GAGGACGTAA AGTTCTAGAT AATAAACCTG GTTCACCTTT  | 5220 |
|    | ACGGAATTGC GAATATGCCA ACGTAACGCA ACAACACCAT AAACAGCCCA AGCATGGAAT  | 5280 |
| 15 | CCCCTAATGGA AAAATGTTGA ACGTAGAGCT TCAGTATAAG CTTCAGTAGT TTTGGGATCT | 5340 |
|    | GCTGTAGGTG GCGTAGCAAA GTGCGCCATC GGTTCAGCTG CACCATAAAA CACCAAACCT  | 5400 |
| 20 | ATCCCATAC CAGCACTAAA CAACATAGCA AACCATGAAA TTGTATTAAA CTCAGGTTTG   | 5460 |
|    | TCATTTGGTT TACCTAGTTT AAGTTTTCCA ATAGGACTAA AAATAAGGAA TATACAGAAG  | 5520 |
|    | AACACGATAA TCGTAGTAAG AATAAGATAA TACCAACCTA ACTTTTCTGT AATCCACATT  | 5580 |
| 25 | TTAATATTAT TGGTAACATA GTTGAATTGT TCAGGTAAAA ATGCCCAAG TAATACGACT   | 5640 |
|    | ATAGCAACAA CAATTGCACT ATAGATGAAG ACTGGTGAAT ACTTCTTTCC ATTTGGATTG  | 5700 |
|    | TCTGGTGAAG AAGAATTCAT AATTAATTAC TCCCTTCAAT TCTATATTTA ATTTTATGTA  | 5760 |
| 30 | GTAGAATAAA AATATTATCT AAACATTTTA TTCAATAACT CACG                   | 5804 |

## (2) INFORMATION FOR SEQ ID NO: 250:

|    |                               |
|----|-------------------------------|
|    | (i) SEQUENCE CHARACTERISTICS: |
| 35 | (A) LENGTH: 400 base pairs    |
|    | (B) TYPE: nucleic acid        |
|    | (C) STRANDEDNESS: double      |
|    | (D) TOPOLOGY: linear          |

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 250:

|    |  |     |
|----|--|-----|
|    | ACCCGCGAAT ATGGTCCATC CTATCGATTT ATTTTAACT GGTTTGACAA TATTTAAATT   | 60  |
|    | TTCAATAATCA TTCTTAGTGA TTTTGACATA TGTTTTCGGT ATGAGCCAGT TAATAAATGG | 120 |
| 45 | AAAGAGAAG ACAAATCCAAT TACTTGCCAA ATCAATCAAT AAATATTCAC TATCGTATTT  | 180 |
|    | GATTATTCGA TATTAGGGT TTTTATTAAT AACTTTAGAT TCGCAAAGCA ATGTCTCCAC   | 240 |
| 50 | ATCCCTTTAA TTTTATGTGT AATACATTTT TCGATACTTC AAAAGACATT CAAATACTAT  | 300 |
|    | CAAGTTACTG TCATCAAAGG TTTTATTAAC TGATATTtTC ATATTTTAAa TCTGAATTTA  | 360 |

55

## (2) INFORMATION FOR SEQ ID NO: 251:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 964 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 251:

|             |            |             |             |            |            |     |
|-------------|------------|-------------|-------------|------------|------------|-----|
| CCAGGGTGGC  | GAAAGCTTTA | AAATTGGAC   | TAATAAAGAT  | GCTGATATTA | ATTCTATGAA | 60  |
| AACAGCAGTT  | TTACAACAAT | TAAAGGAGA   | ATAACATATG  | CTTACTGGcA | AACAAAAAAG | 120 |
| ATACTTAAAG  | AGTTTAGCAC | ACAATATTGA  | TCCGATTTTT  | CAAATTGGAA | AAGGCGGTAT | 180 |
| CAACGAAAAAT | ATGATTAAAC | AAATAGATGA  | TACGTTAGAA  | AACAGAGAAT | TGATTAAAGT | 240 |
| ACATGTACTA  | CAAAATAACT | TTGATGATAA  | AAAAGAATTA  | GCTGAAACAT | TAAGCGAAGC | 300 |
| TACTCATAGT  | GAATTAGTGC | AAGTGATTGG  | ATCTATGATA  | GTGATTTATA | GAGAACTTAA | 360 |
| AGATAATAAA  | GAAATTGAAT | TGCCATAATA  | ATGAAAAAGA  | TATyACTTTA | CGGCGGTCAG | 420 |
| TTTAACCCCTA | TCCATACCTG | ACATATGATA  | GTAGCTAGCG  | AAGTATTTCA | TGAATTACAG | 480 |
| CCAGATGAAT  | TTTTATTTTT | ACCTAGTTTT  | ATGTCCTCCAT | TGAAAAAGCA | CCATGATTTT | 540 |
| ATAGACGTTT  | AGCACAGATT | AACAATGATA  | CAGATGATTA  | TCGACGAGCT | TGGTTTTTGA | 600 |
| GATATTTGTG  | ACGATGAAAT | TAAACGTGGT  | GGTCAAAGTT  | ATACCTATGA | CACGATCAAG | 660 |
| GCATTCAAGG  | AGCAACACAA | AGACAGTGAG  | TTGTACTTTG  | TTATTGGGAC | GGATCAGTAT | 720 |
| AACCAACTAG  | AGAAATGGTA | TCAAATTGAA  | TACTTAAAAAG | AAATGGTTAC | TTTTGTAGTT | 780 |
| GTAATTCGAG  | ACAAAAATAG | TCAAATGTTT  | GAAAAAGCTA  | TGATTGCAAT | TCAGATACCT | 840 |
| AGGGATGATA  | TAGTTTCGAC | AATGATTCTGA | CAAAGAGTTA  | GTGAAGGGAA | ATCTATCCAA | 900 |
| GTTCTTGTTC  | CTAAATCCGT | TGAAAACATAT | ATTAAGGGGG  | AAGGATTATA | TGAACATTGA | 960 |
| AAAA        |            |             |             |            |            | 964 |

## (2) INFORMATION FOR SEQ ID NO: 252:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1193 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 252:



EP 0 786 519 A2

|    |  |      |
|----|--|------|
|    | TTGTCATGAA ATAAATGGGA TGAATATCAC GACTAGAAGT AATGTTACGA ACAGGAGCGT  | 120  |
|    | ATAAACTAGA GACGCTAAAT TCGACATAGT ATGTCGCTCA ATTATGCGTG ATGATGAATT  | 180  |
| 5  | TAAAGTATGT GCGTTGGAAC TGTCGGATTT TTGTCATAA TGTITTCAT ATTGCGCCAT    | 240  |
|    | GATGAATAGT GTAAAAATAA ATAAAACAAT AAGAGATATA ATGCCCATAA TCAAAAGTAT  | 300  |
|    | TTGTTTAGAG CCTTTCATTA TTTCACATCC TTTCTAAAAT ATATTTGTAA CTAAATTTAA  | 360  |
| 10 | AATAGTTATT TTGTAAATC TAAACCTTTT TCATCGCGAA AACAAATAAA TAGGTCGCGG   | 420  |
|    | TATTAATTAT TATATTATTA CCGCTTAATA TGAATAATAC ATGAAATTA ATTTTCTAAT   | 480  |
|    | ATACTTTTGA AAAATTATTA CAAATTAGCC CCTTCAAAAC GCGAAAACAT AAGGATTCTA  | 540  |
| 15 | GTTCAAAAG GCGTGATAAG CATAAATGA AATGTAATAT TTCGATGTAT AAAATTTTAA    | 600  |
|    | ATTAGCTAAA AATCATCGCA TTAATTTTTT GAGCTACAT ATCAAAATTC GGACATTTTA   | 660  |
|    | ACGACACATA TAAITTAATT TTAGGTTTCA TACCAGAAGG ACGTAAAGCG ATAAATCCTT  | 720  |
| 20 | CGTCAAATAA GACACGAATA ACATTGTGATT TAGGAGAATT AATCTGCGAC GTTGATCTCT | 780  |
|    | TATCCAAATG ATAAACCTCG CTAGTTAAAT AATCTTCAAT TGCTTTCACT TTGAGTCCTT  | 840  |
| 25 | GAATCTCTTG CCGTGGATTT GAACGGAATT TGCTCATTAT TGCATTAAIT TTCTTTTCC   | 900  |
|    | CTTCAAATCC TTCTAGCGTA TGCGGAATAA TGTATCCTCA TGTCTACCAA CAGTTTGATA  | 960  |
|    | AATCTGTTCT AATTCATCTT TCAATGTTTT GCCATATAAT TTAACTCAG AAGCGTATTT   | 1020 |
| 30 | TATAATGAGT GGCAGCAATT GTACGSCATC TTTATCAGT ACAAAGGCT CTGATAGAAA    | 1080 |
|    | ACCGTAACTC TCTTCAAATG CGAAATCAT ATTGATGAT CATCCAGTTG TCTTATTTCC    | 1140 |
|    | TGAGCAATAA ATTTAAGCCC GTCAGCACCT CTTTGGTATT CAACATTATT ATA         | 1193 |

(2) INFORMATION FOR SEQ ID NO: 253:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1098 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 253:

|    |   |     |
|----|---|-----|
| 45 | TATCAGGATG ATTATGnCGG GGTTTTATTA AGTCTGAGCG TATATnCTTT TTATCTCTCT | 60  |
|    | AAATGGTTTG GAATTCAGG AGAAGACGAT GAAATGGTAT CCAAAAACCA AAGAAATATAT | 120 |
| 50 | nGCATTAAGT CTGAGGATGA TAGTGGGCTG GCAATTCGTT CATTAATTTT GCATAAAGAT | 180 |
|    | GAACCTATGT ATTTAAAAAA ACGTACATGT GTACCTACTT TGTTAATTAA TGGGGAACAT | 240 |

55

AAAAAAATCT TCGAACATTC AGGACATGCA CGCATATTG AAGAACCAGA AGCATTATG 360  
 AATTATTATT TAAATTTTTT AAAAAGCGTA TCATAATATG TGATATATAA ACCTAGGGCA 420  
 5 TAAAGTCCTT AGGCAATGTG AAAAAGCTGA TTACTATTCA TTATTTGATA GAAATCAGCT 480  
 TTTTTTGAAA TGTATTGTAT ATATACTGCT CGTTATGCGG CTATCTTCCT TATATTAAGT 540  
 GCCATTAGTG CAAAACCTCT TAACAATTAG GTAAAAAGAG CATAAAAAAA GGAAGTTTAA 600  
 10 TAGAATGTAT CATCTATCAA ACTTCACCAA ATTGCGCTAA ACAAAATTAT AGTTCAATTT 660  
 CGTTGTTTGC TTCAGTGATT CGTTTATTIA CTCGACTCAA TAATGATTGG ATTTTITTTAC 720  
 GTTGTGTGTC ATTAACAAGA ATTAATACAG TTCTTTCATC ATGCTCATTG CGTTTTTIAT 780  
 15 CGAAGTAATC TTCTTGAGAT AAAATTTTAA CTGCTTTAAC AACTTGTTGGT TGTITGTAGT 840  
 TTAATGATT AATAATATCT TTAAGATAGT ATTCTTCTC TTTGTTTTCG CTGATGTATG 900  
 20 TCAATACAGC GAATCTTCA AAGCTAATTG AHAATTCCTT TTAATTTAAA CTTTITTAAT 960  
 TGTCAGCATA AGTGACCATT GATAACAATC CAAAGCAATC ATTGATTTTT GTAATTGCCA 1020  
 TGTITAAAAC CTCCTATTTT GATGCATCTT GCTCGATACA TTTGCCCOGA TAATATANTG 1080  
 25 TATCTAATCT TTATGnAT 1098

(2) INFORMATION FOR SEQ ID NO: 254:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 2881 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 254:

CCAGGTAAAA TTGTGCAATC ATTTGACGCA TTAATGGACG CCTTGGACAA TGAAGATTAT 60  
 GAAGGAGAAA AAGTCATTCC ATTCTTAGAT AAACATTTTA AATATCAAGA TGGCCGATCA 120  
 40 AGTGAGCGTT TAGTCAGAAA TTTATTGGT AGCTAAGTTT ATATAGTAGT CAAAGTGGGA 180  
 GAGGTATAAT GATGAAATTT TCAGTAATAG TTCCAACATC CAATTCAGAA AAGTATATAA 240  
 45 CAGAATTACT TAATAGCCTT GCGAAACAG ATTTCCGAA AACTGAATTT GAAGTGGTTG 300  
 TAGTTGATGA CTGTTCAACA GATCAAAAGT TACAAATAGT TGAAAAGTAT CGCAATAAAT 360  
 TGAAC TTGAA AGTAAGTCAA CTCGAAACAA ATTCTGGTGG TCCAGGTAAA CCTAGAAATG 420  
 50 kGGCGTTAAA ACAAGCAGAA GGTGAATTTG TATTATTTGT GGACTCCGAT GACTATATAA 480  
 ACAAGAGAC TTTAAAGGAT GCAGCAGCAT TTATTGATGA ACATCACTCA GATGTCTTAT 540

|    |  |      |
|----|--|------|
|    | CACCTGAAGT TACTTGTGTA AATTCAAGAA TTATCTATAC TTAAAGCCCG ACTAAAACT   | 660  |
|    | ATAGAACAGC ATTACTAAAA GATAATGACA TTTATTTTCC AGAAGAATTA AAGAGTGCAG  | 720  |
| 5  | AAGATCAATT ATTTACAATG AAAGCATATT TAAATGCAAA TCGAATCAGT GTGTTAAGTG  | 780  |
|    | ATAAAGCGTA TTATTATGCT ACAAGCGTG AAGGTGAACA TATGAGTAGT GCGTATGTTT   | 840  |
|    | CACCTGAAGA CTTTATGAA GTCATGAGAT TGATTGCTGT AGAAATATTA AATGCAGATT   | 900  |
| 10 | TAGAAGAAGC CCATAAAAT CAATCTTAG CAGAATTTT AAATCGTCAT TTAGTGTTTT     | 960  |
|    | CTCGTACGAA TGGCTTCTCA CTTAAAGTTA AACTAGAAGA TCAACCAAA TGGATTAAATG  | 1020 |
|    | CTCTAGGAGA CTTTATACAA GCAGTTCCAG AACGTGTAGA TGCAATGGTG ATGAGTAAAT  | 1080 |
| 15 | TACGACCATT GTTGCACTAC GCGAGAGCGA AAGATATAGA CAACTATAGA ACTGTGGAAG  | 1140 |
|    | AAAGTTACCG TCAAGGTCAA TACTACCGTT TTGATATTGT AGATGGTAAA TTAACCATTC  | 1200 |
|    | AATTCATGA AGCGGAACCA TACTTTAAAG GCATTGATAT CGCTAAGCCA AAGTGAAAA    | 1260 |
| 20 | TGACAGCATT TAAATTTGAT AATCATAAAA TTGTTACAGA GCTAACGTTA AATGAATTTA  | 1320 |
|    | TGATTGGCGA AGGACATTAT GATGTGAGAC TTAATTTACA TTCAACGAAA AAGAAGCACA  | 1380 |
| 25 | CAATGTATGT ACCTTTAAGT GTCAATGCGA ATAAACAATA TCGTTTTAAC ATTATGTTAG  | 1440 |
|    | AAGATATTA AGCGTATTTA CCTAAAGAAA AAATTTGGGA TGTTTTCTTA GAAGTCCAAA   | 1500 |
|    | TAGGTACGGA AGTATTTGAA GTGCGTGTG GTAATCAACG TAATAAATAT GCATATACTG   | 1560 |
| 30 | CAGAAACAAG TGCATTAATT CATTGAATA ATGATTTTTA TAGATTAACA CGGTATTCCA   | 1620 |
|    | CAAAAGACTT TAATAACATT TCGTTATACT TTACAGCTAT TACATTAACG GATTCAATCT  | 1680 |
|    | CATTGAAGTT AAAAGGTAAA AACAAAATCA TTTTAACTGG TCTGGATCGT GGTATGTAT   | 1740 |
| 35 | TTGAAGAAGG TATGGCTAGT GTCGTACTAA AAGACGACAT GGTGATGGGA ATGTTAAGCC  | 1800 |
|    | AAACATCAGA AAACGAAGTG CAAATCTTA CTTAGCAAAG ATATTAAAAA GCGAGACTTC   | 1860 |
|    | AAAAATATTG TTAAGTTAAA CACTGCACAT ATCACTTATC CACTAAATAA ATAATAAATG  | 1920 |
| 40 | CCCTCAAAATC ATTGTGAGCC AACATGATTT GAGGGCTTTA TTTTGCTGTT TATGACATGA | 1980 |
|    | TTATGACATT TCCCTGATTT TCATTTTCAT ATACATTAAA TTGTATACAC TGGAAATGAG  | 2040 |
|    | GAGGTATCT ATATGATAA ATAAAAATGA CATAGTAGCA GATGTAGTAA CTGATTATCC    | 2100 |
| 45 | GAAGCAGCG GATATTTTTA GAAGTGTGGG AATAGATTTT TGTGTGGCG GACAAGTAAAG   | 2160 |
|    | TATAGAAGCA GCAGCCTTAG AAAAGAAAAA TGTAGATTGG AACGAATTAT TACAGCGTCT  | 2220 |
| 50 | CAACGACGTT GAACAAACGA ATACACCAGG TTGTTTAAAT CCTAAATTTT TAAATGTTTC  | 2280 |
|    | ATCACTTATT CAATATATT C AATCAGCATA TCATGAACCT CTAAGAGAAG AATTTAAAAA | 2340 |

TGAGTTAAAA GAAACATACG ATACATTTAA AAATGGCATG TTAGAGCATA TGCAAAAAAGA 2460  
 AGACGATGTC GATTTTCCAA AACTCATTAA ATATGAGCAA GGTGAGGTAG TAGACGATAT 2520  
 5 TAATACTGTG ATAGATGATT TAGTTTCAGA CCACATTGCA ACGGAGAAAT TGTTAGTAAA 2580  
 AATGAGCGAA TTAACATCTA GTTATGAACC TCCGATAGAA GCGTGTGGTA CTTGGCGACT 2640  
 TGTTTATCAG AGATTAAAAAG CACTTGAAGT GTTAACACAT GAACAGGTAC ATTTAGAGAA 2700  
 10 TCACGTATTTA TTTAAAAAAG TATCATAAAT AACGCGATTA GAAACTGTGT GCAAAAAATAA 2760  
 GTCCAGCAGT TTTTCGCTAT GTATAAAGT CATAATAGTG ACATAAACAG CATTATTTGA 2820  
 AAAGAANAAT GGTCAACTTA GCATAAAAAAT TGATATGAAN ATTTAATGGT ATAGATAATT 2880  
 15 A 2881

## (2) INFORMATION FOR SEQ ID NO: 255:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1056 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 255:  
 ACOGTCGAAT ATCGCTTGTG ATTTACAATT TGTGTATTAA GATGCTCAAC TAATTGGGT 60  
 30 ACATATTTCCG AATTTAGATT TGCAAGTACA ACAATTCAT AATTTTGT TTGGATTAGT 120  
 AAAATAAATG ATGAAAAGTT ATCTAGCGTT CCGTAATGAA AACTAAATG TTCATCATT 180  
 TTGGTAAACC AGCCGAAGC ATATGCATTG GCATTAGTGT CACCAATTGT TGAAGATAAA 240  
 35 TTTTATGTG ATTGTTGAAC TAATGATTGT TATTATCAG GTGGATTAG TTGGAATTTT 300  
 ATCCGAATGTT CCAAACTTC AGTTGATGTC ATCATATATG CTGATGGTGT ATCCCAAAGG 360  
 TTAATTCAG GTTTAGAGAC GACAGGTGTC GAACCTTGTA ATTCATAGCC AATAGCATCA 420  
 40 TGTTTTGATT TGTAATTGGT TTGTTGAAT GATGTATGTG TCATATGCAA AGGCTTGAGC 480  
 CATGAATTTG TAATATATTT TGTATAGGAT TGCTTCGTAA CGTTTGGAT AATTAAACCT 540  
 45 AATAATCAT AGTTCATATT TGAGTATTCA AATTCCTC CGGGCTTATG ATGTAATTCA 600  
 TCACCCATAA TTGCATGGGT TACATATTT AAACGATTAT TTTGCTGT CACAGAATCT 660  
 TCGCTTGTA TATCACTAGG TATACCACTT GTTTGAGCCA AAAGTTGCTT AATCGTAATA 720  
 50 GTTCATTTT GACCATTATA GTTCATTTTA AAATGAGGCA CATGTTTGA TACGGCATCA 780  
 TTTAAGTTTA ATCGACCTTC TTGAGCTAAT TTTAAATTTG CAAGACCTGT GAAAGCTTTC 840

5  
 10  
 15  
 20  
 25  
 30  
 35  
 40  
 45  
 50  
 55

TGATAACCAT AACCTTTATT TAAAAAACT TTGCCATTTT TTACTAyTAA AATTGATGCT 960  
 CCAGGAATGT GTCCCTTTTG TAAATCATGC TCGATAATTG TATCTATTTG TTGTTGCGAA 1020  
 TCATTGGTTA ACCGTGTCTT CGTATTGCTA TTTAAT 1056

(2) INFORMATION FOR SEQ ID NO: 256:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1277 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 256:

ATGCCaCACT TATTGGTGGC AGGATCGACG GGTAGTGGTA AATCTGTTTG TATAAATGGT 60  
 ATTATTACAA GTATTTTATT AAATGCTAAG CCGCATGAAG TTAACTTAT GTTAATCGAT 120  
 CCGAAAATGG TTGAACATAA TGTTTATAAC GGAATCCAC ACTTATTAAT TCCGGTTGTT 180  
 ACAAACTCCTC ATAAAGCTGC TCAAGCTTTA GAAAAATTG TAGCTGAGAT GGAAAGACGT 240  
 TATGATTAT TCCAACATTC ATCACTAGA AACATTAAAG GTTATAACGA ATTAATCCGT 300  
 AAGCAAAATC AAGAATTAGA TGAGAAGCAA CCAGAATTAC CTTATATCGT TGTATTGTGA 360  
 GATGAGCTTG CAGATTTAAT GATGGTAGCT GGTAAAGAAG TTGAAAATGC GATTCAACGT 420  
 ATTACACAAA TGGCAGGTGC AGCAGGTATA CATTTAATTG TAGCGACACA AAGACCTTCT 480  
 GTGGATGTAA TTACAGGTAT CATTAAAAAT AATATTCCAT CTAGAATAGC TTTTGCTGTG 540  
 AGTTCTCAAA CAGATTCAAG AACTATTATT GGTACTGGCG GCGCAGAAAA GThACTTGGT 600  
 AAAGGTGACA TGTTATACGT TGGAAATGGT GACTCATCAC AAACACGTAT TCAAGGGGCG 660  
 TTTTAAAGTG ACCAAGAGGT GCAAGATGTT GTAAATTATG TAGTAGAACA ACAACAGGCA 720  
 AATTATGTAA AAGAAATGGA ACCAGATGCA CCAGTGGATA AATCGGAAAT GAAAAGTGAA 780  
 GATGCTTTAT ATGaTGAAGC GTATTTGTTT GTTGTGaaC AACAAAGGC aAGTACATCA 840  
 TTGTTACAAC GCCAATTaG AATTGGtTAT AATAGAGCAT CTAGGTTGAT GGATGATTTA 900  
 GAACGCAATC AGGTAAATCG TCCACAAAAA GGAAGCAAGC CTAGACAAGT TTTAATAGAT 960  
 CTTAATAATG ACGAGGTGTA AAAAAATGTC AGAAATGAAT GCGGTATATA ACGTTAAACA 1020  
 ATaCATTttA AATTtgATTA AGCAAAATAA ATTGGAATAT GGTGACCAAC TTCCAAGTAA 1080  
 TTTATCAATT GCCAGAGAAT TAAATGTAAA AACCAGCAT GTTTATGAAG CAATTcAGCa 1140  
 TTGATTACTG AACAGTCAT TAAAGATaT TTGAAGAGG CACAAGTGTT AAGTCACTGC 1200

GrtTTGAATG CGGAAC

1277

(2) INFORMATION FOR SEQ ID NO: 257:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 3557 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 257:

|             |            |            |             |            |            |      |
|-------------|------------|------------|-------------|------------|------------|------|
| TACCGTCGTT  | TTATGyGTCA | AATTTTACAG | TAAATTTTGC  | TTCATCAAAA | GAAATAACCT | 60   |
| TTAACAAAGTA | TAGTAATTTT | ACATTTTACA | ATGTTACAAA  | ATATAATTTT | TTATAATTAG | 120  |
| TTAAATCAC   | TAAACGCTT  | TTATACACTA | TCAAATCAGC  | ATTTATAAAA | ATATGAACCG | 180  |
| ATATCCTAAA  | ATGTTAATAA | TATTACAAGA | TAATAACArA  | CCACACAAAG | CTACTTATTT | 240  |
| TTGATAATAT  | GGAAATCGTA | ATATAAAACn | AAAACTTAAT  | TTACTATATA | AAITGTCTTA | 300  |
| ATAATTTTAA  | AAAGTAGTAA | AACATAATTT | TAAAGAGGAG  | TCCCTTTGAA | AAAATTAGCA | 360  |
| TTTGCAATAA  | CAGCAACATC | TGGTGCAGCT | GCATTTTTAA  | CGCATCATGA | TGCACAAGCT | 420  |
| TCTACACAAC  | ATACAGTACA | ATCTGGTGAA | TCATTATGGA  | GTATTGCTCA | AAATACAAAC | 480  |
| ACTTCAGTAG  | AGAGTATTAA | ACAAAATAAC | CAATTAGATA  | ACAACTTGGT | ATTCCCTGGT | 540  |
| CAAGTTATCT  | CAGTAGGTGG | AAGTGATGCA | CAAAATACGT  | CAAAACCTTC | TCCACAAGCT | 600  |
| GGTTCAGCAT  | CATCTCATAC | TGTACAAGCT | GGTGAATCAT  | TAAATATCAT | TGCTAGCAGA | 660  |
| TATGGTGTIT  | CAGTTGATCA | ATTAAATGCA | GCCAATAACT  | TACGTGGTTA | TTTAATTATG | 720  |
| CCTAACCAAA  | CATTACAAAT | TCCTAATGGT | GGATCAGGTG  | GTACAACACC | AACAGCTACA | 780  |
| ACAGGTAGCA  | ATGGCAATGC | ATCATCTTTT | AATCACCAAA  | ATTTATACAC | TGCTGGTCAA | 840  |
| TGTACATGGT  | ACGTATTTGA | CCGTCGTGCT | CAAGCTGGTA  | GTCCAATTAG | CACATATTGG | 900  |
| TCAGACGCTA  | AGTATTGGGC | TGGTAACGCA | GCTAATGATG  | GTTACCAAGT | AAACAACACA | 960  |
| CCATCAGTTG  | GTTCAATTAT | GCAAAACACA | CCTGGTCCAT  | ATGGTCATGT | TGCTTATGTT | 1020 |
| GAACGTGTCA  | ATGGTGATGG | TAGTATCTTG | ATTTCGTAAA  | TGAATTACAC | ATATGGTCCA | 1080 |
| TACAATATGA  | ACTACCGTAC | AATCCAGCT  | TCAGAAGTTT  | CTAGCTATGC | ATTCATCCAT | 1140 |
| TAATTAATAA  | AATTGTACTG | ATATATACTA | GCAATTACACA | TCATGTGAGA | TTGCTAGTTT | 1200 |
| TTTATTTTTG  | AAAAAAATTT | TCATTTTGGT | ACAAAAAATT  | ATCTCACCTC | TCCTATCAT  | 1260 |
| ACATATTTAT  | ATTTTGTATG | AATGGTAGTT | AGGTAATAAT  | TAACAACCTA | CCTATTTGAT | 1320 |

EP 0 786 519 A2

|    |  |      |
|----|--|------|
|    | ATTAAATTG TTATACCACT ATTTTACGCT TTTTCGTCTA CATATACAAA TTTATATTAA   | 1440 |
|    | ATAAGGCCA ATACAATTTA GGTAAATTAA ACAAGTGAT AACTATTTTA TTATTCCTTC    | 1500 |
| 5  | ATTGAAGAAT ATAAACTATT AAATCATTAT TTTGCTCTTA CATATATTTT AATGACCTTA  | 1560 |
|    | CTGATTATGT TCCATGGAAT ACATTTATAA TATAGCCCTC TAATTAATAT GCYTTGCTCT  | 1620 |
|    | GGTCATTCTA CGTAAATTCG ATAAAATATG TTATCTACTT ACATAAATAT CTGACTTCA   | 1680 |
| 10 | ATACCACCAT ATGTTTGTGA TACTGAAGTT CAGTTTAGTT TTATTTTCAA TTAGAAAAAT  | 1740 |
|    | AAGTTAAGTA TATAGAATAG TAAACCTGCT AACAACTGCT AAATAGGTAA TGTAATCACC  | 1800 |
|    | CATGTAATGA TCATTCGTTG CGCAGTGCTC CATTTTACAC CTTTAGCTCG GTTAGAAGCA  | 1860 |
| 15 | CCAACACCTA AGATTGATGA TGACACAACG TGAGTTGTTG ATAATGGGAA ATGTAGCGAT  | 1920 |
|    | GATGCAACAA AAATTGTATA TGCAGATGAT AAATCGCCCG CAGCACCATT TGCTGGACGT  | 1980 |
|    | ATTTTCATAA TATTACCACC TACAGTTTGT ATAATTTTCC AGCCACCACAT TGCAGTACCA | 2040 |
| 20 | AGCCCCATTG CTGTGCACA GGCAAAATTT ACCCATAACT GTGGTTCAAC ACTGCCATCA   | 2100 |
|    | TTCTGTACAT TAGCGACAAT CAATGCCAAC GTAATAATAC CCATTGATTT TTGCGCATCA  | 2160 |
| 25 | TTCGTACCGT GAGAGAATGA TTGTAACGCT GCTGTGAAAA TTGGAAAAA TCTAAAGTTA   | 2220 |
|    | CGATTGCTCT TTGTTAAATT TGCATTTTAA AAGATAACTT TAAAAATTGA ATACATCAAG  | 2280 |
|    | AAACCAACAC AAAATGCGAT AATCGGTGAA ACGATTAATA CAATAATAAT TTTTGTGAAA  | 2340 |
| 30 | CCTTGTTAAT GTAACACTCC AAATGAGCCT TCAGATGCGA TTGCTGCACC CGCAATTGAA  | 2400 |
|    | CCTATAAGTG CATGTAAGA CGAACTTGGA ATTCCGTAAA ACCAAGTAGC TAAATTCCAA   | 2460 |
|    | ATAATAGCCG CAAGTATTGC AGCTAACACA ACACTAATC CATTTTCCAA TTTAAATGGA   | 2520 |
| 35 | TCGACATGTT CTTTAGTAAT GTGCGCTGCA ACGCCCGTAA ATGTTAAAGC ACCTATAAAG  | 2580 |
|    | TTCACTCACTG CTGCCATTAA AATTGCCGTT TTAGGgTIAA CGCTCTAGTA GATACAGCAG | 2640 |
| 40 | TAGCTACTGC ATTGGCTGTA TCATGGAAAT CcATTGTAA AGTCAAAATAT CAGCGAGAAA  | 2700 |
|    | ATAACTACAG CTATAGTGAC GATGATTATA TATGACATAA ATATATACTC CCCTTAGCTA  | 2760 |
|    | TTTTTCATAA TAATAGTTTC AAAATTATTT GCTACGATTT GACATTTATC AGCGATTTC   | 2820 |
| 45 | TCCATGCTTT CATAAATATC TTTTATTTTA ATTAAGTGA TTGATCTGTT TTGCTATTG    | 2880 |
|    | AAAATATGTT TAATTGACTG TCTTAAAAATA CCATCAGAT TTGTTTCAA TTCTTTAATA   | 2940 |
|    | TAAATTGAAT GAATACGCAT ATGTGATAAT TTTTATCGA CTAATAAGCC GACAGCAAGT   | 3000 |
| 50 | TTCAATTCTG CAAGTCTTTT TTGAATGTTA TCAACAACT CAGCCATATA TTCATCTGTG   | 3060 |
|    | TATTGCGATTG AATACATTTT AAACATGCTT GCCGTTTCTT CAATTGCATC TAAACATCA  | 3120 |

55

EP 0 786 519 A2

TTTAAATCAG TAATTACTTG ATGTACTAAT ECGcACCATG TGACTCATAA GTTTTAAATGT 3240  
 TGCTGAGTA TGCTTTTAAA TCTAAATGTG TATTGAAATC CATTTTACCG AATTCAATAG 3300  
 5 CAGCAGGATC CAGATTGAAA ACCATCTCTT CTAATTGAAC CATAACTTA TCTTTTCTCT 3360  
 TACTAAACAT TTAAAATCCT CCAATTAAAGC GATTGTCAAC AATCACATTC AGTTATAATT 3420  
 TGTTTCAAAAT TAAGACAAGT GAATTIACAA ACTAATGATA CAAATTTGTT ATTATCAATC 3480  
 10 GTCAGTATAA TTTTAGTGTA CTGATATTAA TTTCAAAAAT GCCTCACAGT AAACAATTTA 3540  
 CTGTATTGTC CCTTATA 3557

(2) INFORMATION FOR SEQ ID NO: 258:

15 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1631 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear  
 20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 258:

AACTATACAT TTGAAAAAAT TCTTCTAGTG AACCTGGGCC ACCAGGAGCC ATGACAAATG 60  
 25 CATCTGCAAG TTCTGCCATT TTATTTTnAC GTTCATGCAT AGAATCAACT AAAATTAAAT 120  
 CAGTTAAAGC TTGGCTTGTG ATTTTCATGT CATCTAACAT TTTAGGCATG ACGCCAATAG 180  
 30 CTTTGCCGCC ATGATCTAAT ACACCATCTT GaATGGCACC CATAATGCCA ATTGACCCTG 240  
 CACCAAAATAC TAATTCATAA CCTTGTTTCAG CAAAATATTT ACCTAAATCG TATGCTTTTT 300  
 GTACATATGA AGGSTCATGA CCTTGTCTTG CACCACAATA AACTGCGATT CGTTTCATGT 360  
 35 TAATCCAGCT CCTTAATTG ATGAATGACT TTTAATAGTG ATTGTTCAAA CACTTTTGA 420  
 TCTTCTCTTG TAAAGGTG GGGACCTTTG TGGCGACCAC CTGTTTTCT AATTGTGCA 480  
 TTCAATATATC GTTATCTAA TAGTGTGTTGA ATATTTTTGG AATTGTATAT CTTCACATTA 540  
 40 TGATGCATGA CAATTAAGAC TTTGTGCACT AATAAACTTG CGAGTCCATA ATCTTGAGTG 600  
 ACTIACGATAT CATCCTTCGT TGATAATTGA ACAATTTGT AATCAACTGC ATCTGGTCCA 660  
 TCATCAACAT ATAATGTTGA TACATGTGGA GGATATAAAT GGTTCGAAAA ATGGCTGAAG 720  
 45 CTCGGAATAA TTGTACAAAA AATGCCTGTC TCAGTTGTTA AATCTATAAT AGAATCAACA 780  
 ACAGGACAAG CATCTCCATC AATAATAATA TGTGTACAAA TTATGCTCTT GTATTGTTTT 840  
 CTTTATTTTT TTGAGAGGCG CTTTGTGGCA CATAATCTTT ATATTTTTTA AATGACTTGA 900  
 50 TGCGTGCTTT ATCAGCTTCT TGTGGCGT TTTGTCTTC TTTGTGCTT TTTCAATAT 960

55



CGCCCTTTTTC CTCAGTTTTC TCATCTAATT TAITAGGTGT TAAGCCTGCT TTTTCTTCGT 1080  
 ATTTTGTGA TTTTTCATA TCCTTAANTAC GTGTGATTC ATTCTTTTCG CGGCCTTTT 1140  
 5 GCTCTTCTTT ATGACGCTTT TCGATATTTT TTTGAAGTAT TTTATTCATT TTATCAGCGT 1200  
 CTTTACGATT TTGTTAGCT AATTTTTCG CTTTTTCTC AATATAGGCA GGATCATGTT 1260  
 CTCTAGCAAA CTTTTTAAGT TCACGTTTAT TTICAAAATC TTGTTTTTTA TCGCCGACAT 1320  
 10 ATTCTTTAAC ATCACTCGCT GTGTACTGA TTGCTGCAGA TGTTTTTGAA GCAACTTTAC 1380  
 TTGTAGCATC TGTAACCTTT TGTACGTCGG GATGTTGTTT GATACGTTTA CGTTCAACAA 1440  
 TTAACGCTAC CAATACAATT GGTAATACAT TAATCATAAA TTTGATGACT TTTTCTTTAT 1500  
 15 CCATAGATCT TGCCTCCaTA ATTACTTTAT TAACTTTTCA TACCCTATGA TACATCAATA 1560  
 TAAACGATGA TAGTAGTGAA TCACTATTAA GTATTTCAGA TGTTTTTAA AagaAGaCCC 1620  
 AATTAAAAA A 1631

(2) INFORMATION FOR SEQ ID NO: 259:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 5645 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 259:

CGAAATCATG ATTTAATGCT TTTTCATATA AGCTTTTCCA ATTAATCTTT CGTCCATGAT 60  
 ATTCTTCAAC TGTGTGAGA TATTGTGCAA TTTTAGTAC TTTAAAGGAG TGTGCTGCAA 120  
 35 CaTTGTGkTC mAAATATTTA AATTTTCCaG GtAATCTTAT AAGTCTTTCC aTATCTGATA 180  
 ATCTTTTAAA ATATTGATGT ACACCCATTT CAATTACCTC CTCCATTAAAT TAATCATAAA 240  
 TTATACTTTC TTTTACATA TCAATCAATT AAATATCATT TAAATATCTT CTTTaTATAA 300  
 40 CTCTGATTAA ATGATACCAA AAAATCCCTCT CAACCTGTTA CTTAAACAGG CTAAGAGGAT 360  
 AGTCTTGCTC TGATATAATTA CTTAGTGGAT GTAATTATAT TTTCTGGAT TTAATAATTGT 420  
 TCTTGAAGAT TTAACATTAA ATCCAGCATA GTTCATTTCa GAAACAGTAA TTGTTCCATT 480  
 45 AGGGTTTACA GATTCAACAA CACCAACATG TCCATATGGA CCAGCAGCTG TTTGGAAAAAT 540  
 AGGCCCAACT TCTGGIGITT TATCTACTTT AAATCCTGCA ACTTTTGCTG CGTAATTCCA 600  
 GTTATTTGCA TTGCCCCATA AACTTCTTAT ACTTCTACCT AATTGTGCAC GACGATCGAA 660  
 50 AGCATAATAT GTGCAGTTTC CATAAGCATA TAAGTTTCCCT CTGTTAGCAA CTGATTATT 720

|    |             |            |             |            |             |             |      |
|----|-------------|------------|-------------|------------|-------------|-------------|------|
|    | TACATTAAC   | GTCTTAGTTA | CTGCTTGCTT  | AGGTGCTTGC | TAACTACTA   | CTTTTITAGA  | 840  |
|    | TGCTTGTTGT  | ACAGGTTGTT | TTACTACCTT  | TTAGCTTGG  | CTTGCTTTTC  | TTACTGTGTA  | 900  |
| 5  | TTTAAACCGT  | TTAGTTTGT  | TCACCTTATT  | TTGAGGCACA | AGTGAAATCA  | CGTCACCAGG  | 960  |
|    | AAAAATTA    | GGTGTACAC  | CAGGATTGTA  | TGAATATAAT | TGATTCAACG  | TTAAGTGATG  | 1020 |
|    | CTCTAAAGCA  | ATCTTATATA | ATGAATCGCC  | AGCAACTACT | GTATAAGTTG  | TCGGTGATTG  | 1080 |
| 10 | CGTTTGTGCT  | TGAACATTG  | ATACATAATT  | ATGTTGAACA | GGTGTITTTA  | CTTGTGTGCC  | 1140 |
|    | ATGTTGTGTG  | GCATGTGCTG | CATTATTTAA  | AGCTAAAAAA | GCTAACACTG  | ACGAAACCGT  | 1200 |
|    | CACGTGAAGA  | GATTTTTTCA | TCTTGCTGTC  | ATTCCTTTGC | TGTTAGTATT  | TTAAGTATGC  | 1260 |
| 15 | AAATACATATA | GCACAATACA | TTTTGTCAAA  | AGCTATTGTT | ATAACGATGT  | AATCAAAATGG | 1320 |
|    | TTAACATAT   | AAAAAGAATA | CAACCTTTTA  | TCATAGTGTA | AAATGTATTG  | ATACCATGTA  | 1380 |
| 20 | ATTGAGAACG  | TTTTCAATAA | TTAATTCAAT  | ACCTTGAAAA | TCGCCATAGG  | TAATATTACT  | 1440 |
|    | AAATGCACAC  | TGCATATGTT | GTTTAAACAA  | ACACAACCTT | TAAAAAATAT  | ATCTAATCTC  | 1500 |
|    | TATCTACCGA  | ATTGACTCTA | AATATTCATA  | AACAAATCAT | ATTCAAAAAT  | CTAATTTACA  | 1560 |
| 25 | ATTTATTTAG  | CTACCTTTAA | AAAAACCAAA  | AACCGACGCC | CTTTAGAGC   | CTCGGTTTTA  | 1620 |
|    | AAATATATAT  | TAATCGTGCG | ACATTGTCTG  | TCTTAAATAT | GATTCGATAA  | ATGGTCCAAT  | 1680 |
|    | GTCCTCATCC  | ATCACTGCAT | CAACCTTACC  | TGTTTCTTCG | TTCTGACGAT  | GATCTTTCAC  | 1740 |
| 30 | CATTGAGTAT  | GGATGGAAAA | CATATGATCT  | AATTGGGCTT | CCCCAGCCGA  | TTCTTTTTTG  | 1800 |
|    | TTGCCACGA   | ATTTCAGCCA | TTTCACGTGC  | CTGCTCTTCC | AATTTTAAAT  | GATATAAATT  | 1860 |
|    | AGACTTTAAC  | ATTTTCATAG | CTGCTTCACG  | GTITTTAATT | TGAGAACGTT  | CATTTTGGTT  | 1920 |
| 35 | ATTAACAACT  | ATACCTGAGG | GGTGGTGGGT  | AATTCGTATT | GCCGATTCAG  | TTTTGTTAAT  | 1980 |
|    | ATGCTGACCA  | CCTGCACCAG | AAGCTCTGAA  | TGTATCAACT | GTAATATCAT  | CCGGATTGAT  | 2040 |
| 40 | TTCAATCTCT  | ATTTCATCAT | TATTAATAATC | TGGAATAACG | TCGCATGATG  | CAATATGATG  | 2100 |
|    | ATGACGACGT  | CCTGATGAAT | CAAAATGAGA  | AATTCGTACT | AGTCGGTGTA  | CACCTTTTTC  | 2160 |
|    | AGCTTTTAAA  | TAACCATAG  | CATTATGCCC  | TTTGATGAGC | AATGTATACAC | TTTTAATCCC  | 2220 |
| 45 | CGCTTCATCC  | CCAGGTAGAT | AATCAACAGT  | TTCAACTTTA | AAGCCTTTCT  | TCTCACAATA  | 2280 |
|    | ACGTTGATAC  | ATTCTAAATA | GCAATATTAGC | CCAATCTTGA | GACTCGGTGC  | CACCTGCACC  | 2340 |
|    | AGGATGTAA   | TCTAGAATTG | CGTTATTGSC  | ATCGTGAGGC | CCATCTAATA  | ATAATTGCAA  | 2400 |
| 50 | TTCGTATTCA  | TCCACTTTAG | CCTTAAATTT  | AATGACCTCT | TGCTCTAAGT  | CTTCTTTCAT  | 2460 |
|    | TTCTTCATCA  | AATTCCTCTT | GTAATAAATC  | CCAAGTAGCA | TCCATGTCAT  | CTACTTCTGC  | 2520 |

|    |   |      |
|----|---|------|
|    | TTGCGCTTTC GTTGGTTAT CCCAAAAATT AGGTTCTGCC ATCATTTCCT CATATTCITG    | 2640 |
|    | AATATTAGIT TCTTTGTCT CTAAGTCAAA GAGACCCCT AATTGTGTT AAATCTTGAT      | 2700 |
| 5  | TATACTTATC TATATTCGT TTGATTTCTG ATAATTCAT AGcATTCGCT CCTATTTATA     | 2760 |
|    | TTTCAATTCA AGTCATTGAT TTGCATCTTT TATAATGCTA AATTTAACA TAATTTTGTT    | 2820 |
|    | AAATAACAAT GTTAAGAAAT ATAAGCACAC TGACAATTAG TTTATGCATT TATTTGTAAA   | 2880 |
| 10 | AATCAGTACA TTTATCATCG ACATATGCGT AAACCGATT TTTAAAACTA AGTACATAAC    | 2940 |
|    | AACGTTTAAC AACTTCTTCA CATTTTTTAA AGTATTTAAC GCTTGTA AAAA TAAAAAGACT | 3000 |
|    | CCTCCCATAA CACAACTAT AGGTGTTTAA TTGGAAGGAG TTATTTTATA TCATTTATTT    | 3060 |
| 15 | TCCATGGCAA TTTTGAATT TTTTACCACT ACCACATGGA CAATCATCGT TAGACCAAC     | 3120 |
|    | TTGATCGCCT TTAACGATTG GTTTCGGTTT CACTTTTTCT TTACCATCTT CAGCTGAAAC   | 3180 |
|    | GTGCTTCGCT TCACCAAAC CTGTTGTTTT TTCACGTCA ATATTATCTT CAACTTGTAC     | 3240 |
| 20 | TACAGATTTT AAAATGAAT TACAAGTATC TTCTTCAATA TTTTGCATCA TGATATCAAA    | 3300 |
|    | TAATTCATGA CCTTCATTTT GATAGTCACG TAATGGATTT TGTGTGTCAT AAGAACGTAA   | 3360 |
| 25 | GTGAATACCT TGACGTAAT GATCCATTGT GTCGATATGA TCAGTCCAAT GCCTATCAAT    | 3420 |
|    | AGAACGAAGT AAAATCATAC GCTCAAAC TC ATCATTTGT TCTTCTAAGA TATCTTTTGT   | 3480 |
|    | ACTTTGATAT GCTGCTTCAA TCTTAGCCCA AACGACTTCG AAAATATCTT CAGCATCTTT   | 3540 |
| 30 | ACCTTTGATA TCATCTCTG TAATGTCACC TTCTTGTAAG AAGATGTCAT TAATGTAGTC    | 3600 |
|    | GATGAATGTT TGATATTCAG GCTCGTCATC TGCTGTATTA ATATAGTAAT TGATACTACG   | 3660 |
|    | TTGTAACTTT GAACGTAGCA TTGCATCTAC AACCTGAGAG CTGTCTTCTT CATCAATAAT   | 3720 |
| 35 | ACTATTCTTT TCGTTATAGA TAAATTCACG TTGTTTACGT AATACTTCAT CGTATTCTAA   | 3780 |
|    | GATACTGTTT CGCGCGTCGA AGTATATTACC TTCTACACGT TTTTGTGCTG ATTCTACAGC  | 3840 |
|    | TCTTGATACC ATTTTGTATT CAATTGGTGT AGAGTCATCT AAACCTAGTC GGCTCATCAT   | 3900 |
| 40 | TTTCTGTAAA CGTTCGAAC CAAAACGAAT CATTAATTCA TCTTGTAATG ATAAATAGAA    | 3960 |
|    | GGCACTATCC CTTTATCAC CTTGACGTCC AGAACGACCA CGTAACCTGGT CATCAATACG   | 4020 |
|    | ACGAGATTCA TGTGCTCTG TACCTATTAC TGCTAAACCG CCTAATTCCT CTACGCCCTC    | 4080 |
| 45 | ACCTAATTTG ATATCTGTAC CACGACCAGC CATGTTAGTG GCAATAGTAA CGGACCTTTT   | 4140 |
|    | TTGTCCAGCG CCTGCAACAA TTTGAGCTTC ACGTTCATGA TTTTTCGAT TTAACACATC    | 4200 |
| 50 | ATGACGGATA CCACGTTTTT TAAGTAAATT TGAATATAT TCAGAAGTCT CAACTGCAAC    | 4260 |
|    | AGTACCTAAT AGCACTGGTT GCCCTGCCTT GTGTTTTTCA ACAACATCTT CTACTACTGC   | 4320 |

|    |  |      |
|----|--|------|
|    | TTTATTGTC GGAATTGAG TTAAGTGCAT GTTATAATA TTCTAAAT CTCTCTTC         | 4440 |
|    | AGTTTTAGCT GTACCTGCA TACCGCAAG TTTATTGTAC ATTCTGAAAT AGTTTTGGAA    | 4500 |
| 5  | TGTAATAGAC GCCATAGTTT TAGATTCAAT TTGAATTGGA ACGCCTTCCT TCGCTTCAAT  | 4560 |
|    | AGCTTGGTGT AAACCTCCG AGAAACGACG GCCTGGCATT GTACGTCCTG TAAATTGATC   | 4620 |
|    | GACAATTAAT ACTTCGCCAT CAACAACCAT ATAGTCTACG TCACGTTGTA ATGTAACGTG  | 4680 |
| 10 | CGCAAGTAAA GCTGTGTTGA TATGACTAAT AACATCAACA TTTGTACAT CATATAAGTT   | 4740 |
|    | TTCAACTTGT AACATACGTT CAGCTTTATC CGCACCTTGT TCTGTAAAT GTACAGCTTT   | 4800 |
|    | CGTTTTTCA TCGTAATTAT AATCTTCGTC CTGTTTTAAC ATTTTCGCAA AAACATTGTC   | 4860 |
| 15 | TTGTGTATAA AGTGACGTTG ACTTTTCAGC TTCACGAGAA ATAATTAATG GCGTACGTC   | 4920 |
|    | CTCGTCGATT AAAATTGAGT CAACCTCATC AATGATTGCA AAATGTAATG GACGATTAC   | 4980 |
| 20 | TCTATCTTCA GAATAATTCA CCAATGTTATC TCGTAAGTAA TCAAAACCTA GCTCATTAT  | 5040 |
|    | AGTACTGTAA GTAATGTCTT GTGCGTATGC TTCAAGTTTT TCTTCTGTGC TCTTACTGTT  | 5100 |
|    | TAAGTTTAAT CCGACAGTCA AACCTAAGAA GTTATATAAC TCAGCCATTT CTTCACTTTG  | 5160 |
| 25 | AACACTTGAT AAGTATTCAT TGACTGTAAT AACGTGAACA CCTCTACCAG CTAATGCAAT  | 5220 |
|    | TAAGTAATGT GGCATTGTG CTGTTAATGT TTTACCTTCA CCTGTTCTCA TCTCAGCGAT   | 5280 |
|    | ATCACCTTTA TGAATTGCAA TACCACCCAT AATTGAACT TTATATGGTG TCATATTGAA   | 5340 |
| 30 | TACAAGTTTA GAGCCTTCTC TAACAAGTGC ATATGCTTCT GGTAAAATTT TATCTAAATA  | 5400 |
|    | ATCATTTTGC TTTTGTACAT TATCAATGTC AGCTAATCT GTTTGGAATT GTTTCTGTTT   | 5460 |
|    | ATTACGAATT TCTTCATCAG TTAATAATTG CGTTTTTCTC TCTAAGCGGA TTACTTTATC  | 5520 |
| 35 | AGCAAGTTTA CCTAAGTGT TAATTTCCTT ATTATTGCCA TCAAGAATTT TTGATAAAAA   | 5580 |
|    | TCCGATTTCG TTCGCTCCTT TAGCTAAAAA ACTGTTTGGC CTACAACAAT ATATCTTATC  | 5640 |
|    | ATTATATAGT AGAAAAATAT ACTTATTTAC TCATTGTGAG AATCAATATA AATATATTTA  | 5700 |
| 40 | TGACATACCTT CATTACACTT CTGTTGTCAA CAAGTTTATC ACTAATAAAT ATATTCTCAA | 5760 |
|    | TACGCAATTA TACTTCTTAA TAAATTATAT TATAAATATT TTACGATTTT CGACTCGGAC  | 5820 |
|    | TATACAATAG ACTGACATAC TATTATTAACT TTAACATTCA AATATATACA TCCATTAACT | 5880 |
| 45 | TTAGCATAGT CACTATGTTT CATTCAACAA ATTACATTAT CGAACTATGA AATAGTCATA  | 5940 |
|    | ATTTGCTTTT GGAGTATAAA AAAGCACTTG TGCAAAAAA CAAGTGCTTT AAACCTAATT   | 6000 |
| 50 | TATTGTTTAC TAGTTTGAAT CAAGCCATAT TTACGCTCTT TACGCGGTA AACGATACTT   | 6060 |
|    | GTTCACATGAC TTTCTCTGTC TGTGAATACA AAGAAGTCAT GACCTAATAG ATTCAATTGT | 6120 |

ATCTCGTTAT CATCGTAAGC GTCATTATCA ACTTGTGTTT CTTCGATTTC TTGTAATTCG 6240  
 GCAACAAACA CTCTTGATC TCCTCGATCA CGGCTCTTAC GATTAATAAG TGTTTTATAT 6300  
 5 TTTCGAACIT GTCTTCAAG TTTATTATTA ATTAATcAA TACCTGCGTA TAAATCATCG 6360  
 TTTGCTCTTT CAGCTCTTAA CGTAACATTT TTCAATGGAA TTGTTACTTC AATTTTAGTA 6420  
 GCTGAATTTG AATAAGTTTT AACTTTAACA TGCGCCACTG CATTGGGTAC GTCAATTAAA 6480  
 10 TAACGTTCCA ACTTACCAAT TTTTTCCKCA ATATAGTTGC GAATAGCATC TGTGATAGTG 6540  
 AGGTATCTC CATGAATTC AAATCTAATC ATAGTAAATC TCTCCTTAAA CCTCTTATn 6600  
 GGnAACTCht TATTATATTT AACATTTTTA CGCCAATCGT GCAAA 6645  
 15

## (2) INFORMATION FOR SEQ ID NO: 260:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 7430 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 260:  
 25 CAGTCCAGC ACATCTATTG GGGATCAACA AACTAGGGAA AATGCTAATT ATCAACGTGA 60  
 AAACGGTGTT GACGAACAGC AACATACTGA AAATTTAACT AAGAACTGC ATAATGATAA 120  
 30 AACAAATATCA GAAGAAAATC ATCGTAAAC AGATGATTGG AATAAGATC AACTAAAGGA 180  
 TGATAAAAAA TCATCGCTTA ATAATAAAAA TATTCAACGT GATACAACAA AAAATAACAA 240  
 TGCTAATCCT AGCGATGTAA ATCAAGGGTT AGAACGCGT ATTATGATG GTAAACAAAG 300  
 35 TAAAGTGGCG TCACAGCAAC AGTCAAAAGA GGCAGATAAT AGTCAAGATT CAAACGCTAA 360  
 TAACATCTTA CCTTCACAAA GTCGAATAAA GGAAGCACCA TCATTAAATA AGTTAGATCA 420  
 AACAACTCAA CGAGAAATTG TTAATGAGAC AGAAATAGAG AAAGTACAA CACAACAAA 480  
 40 TAATCAAGCG AATGATAAAA TTAATACTA CAATTTTAAC AATGAACAAG AAGTGAAACC 540  
 TCAAAAAGAC GAAAAACAC TATCAGTTTC AGATTTAAAA AACATCAAA AATCACCAGT 600  
 AGAACCAACA AAGGACAATG ACAAGAAAAA TGGATTAAAT TTATTAATAA GTAGTGCAGT 660  
 45 AGCAACGTTA CCAACAAAG GGAACAAAGGA ACTTACTGCA AAAGCGAAAG ATGATCAAAC 720  
 GAATAAAGTT GCCAACAAAG GGCAGTATAA AAATCAGGAT CCTATCGTTT TAGTGCATGG 780  
 50 TTTCAATGGG TTTACAGATG ATATTAATCC TTCAGTGTTA GCTCATTATT GGGCGGTAA 840  
 TAAATGAAC ATTGCCAAG ATTTAGAAGA AAATGGTTAC AAAGCTTATG AAGCAAGTAT 900



|    |            |            |             |            |            |             |      |
|----|------------|------------|-------------|------------|------------|-------------|------|
|    | TCTAAGAAAT | TTAATACAG  | TTCTAGTTA   | TAACCCATAT | ATGCACCTAA | GAAGAAATAA  | 2820 |
|    | AAATCCATC  | CGAATATTAT | AGTATTTTCA  | CTTAATGGAT | AATAGTGTAG | CACGGTATCG  | 2880 |
| 5  | TGAAACGCTG | TGTTGTTGCT | AAAGTAATAT  | AAAAATGATT | GCTGTAAAAT | AAAAGATAAC  | 2940 |
|    | AATAATAATA | TTTTACTGTT | GAATAGGTTA  | TAGTTAATTT | TAAAAATGAT | ATAACTCAAA  | 3000 |
|    | ATAAGAATT  | GCATGATAAC | AACGATAAAA  | TAGCCATACC | ATTGACCTAA | TAGGACATTT  | 3060 |
| 10 | TCAATGAATT | GTTTATTGAA | ACTTGAATCT  | GTTAATAATG | ATTCACCTAA | ACTGTAAAAAC | 3120 |
|    | AATCCCATTA | ATATGTAAGG | AATAAGTATA  | TATTTTACGC | GTGTAGTTAA | GTATCTATAG  | 3180 |
|    | GTGACTTTTT | GGTAATTCOA | GGTTGTCACT  | AACCTGTACA | AGATAATAAA | GCAAGGTGTA  | 3240 |
| 15 | CCAAAAATCA | CAATATTACG | AATGTAAAAT  | TGTAACACTA | AGGATCCACC | CTCCATATTT  | 3300 |
|    | TCATGTTTTA | AAGTAATTG  | TGTAAGTAAA  | TGTGTGATAA | TAATAATTGC | ACATATAATA  | 3360 |
|    | GCACGTAAT  | ATACGAGTTC | AAGTCTAATC  | TTTTTCATGG | AATCCGTCCC | ATCTCTTAAT  | 3420 |
|    | TAAATGCTCA | AAAGCATCAT | CACATAATTAA | TATTTCTAGG | ATGTAATAAT | CATTGGAGTT  | 3480 |
|    | CGGAGTGACT | GCTTTTTCTT | CTAATGAAAA  | ACCGTATTTT | AACCCAGCTT | TTTTGATTAC  | 3540 |
| 25 | CGGTAATTTA | TGCTCATTTA | TCAAGCCATA  | AGGATAGGCT | ATAGTTTTCT | CGGACTTTTT  | 3600 |
|    | AAAGTTTTTA | GTTAGATATT | TTTCACTTTT  | GTTTAAATCT | TTTATGATTG | GATCTTCAGA  | 3660 |
|    | AGCTTTCATT | AATTTTGACT | TATTATTTTT  | AGATAAGTTA | TGCAATCGT  | GGGTATGTGT  | 3720 |
| 30 | TTCAAATTC  | CATAACCCAG | TTTTATACAT  | TTCTTTTAGT | TCTTTTTTAC | TAATCATATC  | 3780 |
|    | GAGGTTGTGA | AAGTTTTCTT | CCCCAACATG  | ACCTGTGATA | ATAAACCCAG | TTGCCGGTAT  | 3840 |
|    | TTTATATTTT | TTTAAGATTG | GATAAGCAAT  | TTCTATAATA | GTTTCATCCA | TATCATCAAA  | 3900 |
| 35 | GTTAATCCAT | ACACTTCGTT | TTGGAAACTT  | ACCTTTTTTC | TTGCAATATA | AAAATTCITT  | 3960 |
|    | CAAGGTTAAA | AATTTAGCAT | CATGTGATTT  | TAGCCATTTT | ATTTGAGATT | CAAATTTGTA  | 4020 |
|    | TTGACTAACA | CTATAATTTT | TAATTTCTTT  | ACTACTAGAA | AAGAAGTAAA | TAAATTTATT  | 4080 |
| 40 | CAGAAAATTC | GCTTTTCTTA | CACGGTGATA  | ATTTAATGCC | AGAGCACTAT | TTTCTTTATA  | 4140 |
|    | TTTCAGTTTT | TTAGGTGAAT | CGTCATCTGC  | ATTTGCAATA | TGATGACCAT | CCAGTGTGCT  | 4200 |
|    | TACAGGCAAT | ATGATCAAGA | TACTCAACAC  | TAAAAATATA | AATTTTCTAT | ACTTCACGAT  | 4260 |
| 45 | TCTCTTCTCT | TCTGCCATTT | TTGAATCAAT  | ATGCTAATTG | TAAAAAATAC | AAAAATGATA  | 4320 |
|    | ATCGCGAAAA | TGCCCATAGT | TTCAAATATA  | TCTAAAAATT | CAGTATTTTC | AATGTTTAAA  | 4380 |
| 50 | GCAACACGTA | TGTATTGAT  | ACTTTCGTCA  | TGAATTTCAA | ATATAGTACC | AATATAAAACG | 4440 |
|    | AGTAGAACAA | CTAAACAATA | TATCCAAAG   | ACACAAGATA | TAGCGATAAG | TGCTGTTTCT  | 4500 |

|    |             |             |             |              |             |             |      |
|----|-------------|-------------|-------------|--------------|-------------|-------------|------|
|    | TTGCGTAACC  | ACCTTTCTTA  | CGTTTTAATG  | CTTTTGGAAA   | TGCGACAAGA  | ACTACTGCTG  | 4620 |
|    | CGTTAATAAT  | CCAGTATACT  | GTCGGATACC  | AACCTACAAA   | TATGAGTCCA  | GCCATATTCT  | 4680 |
| 5  | TTTTCTCGTA  | GGGACTATCA  | ATAAAGAGTG  | CGACTGTAAA   | TTGAATAACG  | TTTATAAAAAG | 4740 |
|    | TCATAGTAAA  | TGATGATAGT  | AGAAATATTG  | AAAACTATA    | TGTCATAAAT  | GTATAGTCTTA | 4800 |
|    | AGAAGTTTGC  | TGTTATGAAC  | AAATAGCCTA  | AATATAGAAG   | CACTATATAT  | ACCCATAAAA  | 4860 |
| 10 | TCGAGATGAT  | TTGCTCAAAC  | ATCAAAATAT  | ATAAAGGAAA   | CCTTTTCGTT  | TTCAATTGTGC | 4920 |
|    | TAAAAAAGTC  | TGATAGTAAT  | ACTTCGTGTC  | CCCCTTGAGC   | CCATCTCAGC  | CGTTGCTTCC  | 4980 |
|    | AAAGACCTCC  | CAATGTTTCT  | GGAAACCAACA | TCCAACACAT   | GGCAAGCGGT  | TCATACCTTAA | 5040 |
| 15 | TAGCATATCC  | AGGTAATGTC  | AAITTTCCAAG | AAACTGCAAT   | ATCTTCGGTA  | ATCATATCAG  | 5100 |
|    | TATCCAGTA   | GCCAAAGTCG  | ACAACTGCAC  | TTTTTTTAAA   | TAGAGTGAAG  | ACACCCGAAA  | 5160 |
| 20 | TAGTATTGAC  | TGCGCCAGCA  | AGTGTCTGAC  | TtCGCTTAAT   | ACAGCCAATT  | AAACTTGCAT  | 5220 |
|    | ATTCTATCGT  | TTGAATTTTA  | CCTAAAATAG  | AACTCTTATT   | TGGAATTCTA  | GGATTACCTG  | 5280 |
|    | TAACCTGACC  | AAGTTTGGGA  | TCATGTTTGA  | AAITTCCTCAAT | CATATAATAT  | GGTGCATCTT  | 5340 |
| 25 | GATCAACGAT  | AGTATCTGCA  | TCCAAGCACA  | TTACATAATC   | ATATGAAGCC  | TGTTTAAATGC | 5400 |
|    | CTGATTGAG   | TGCGTTGGCT  | TTACCTCTGT  | TTTCTTGTA    | ATCGACGAAA  | ATAAAGTCAT  | 5460 |
|    | TATTTTCTTT  | GATTTTATAG  | ATGAGTCTG   | CTGTATTATC   | TGAACCTCCA  | TCATTAAATGA | 5520 |
| 30 | TAATAATTTT  | TTTCTTCTG   | TATTTGAGTG  | CAAGAACATT   | AGACAACGTA  | TCTTCAATCG  | 5580 |
|    | TTTCACTTTT  | GTTATAACAG  | GCAAGTAAAA  | ATGTAATGCC   | TTCTAATTCA  | TCCACATTTA  | 5640 |
|    | TGTCAGGCTT  | CTGTGTTCAAT | GAATATCTAA  | TTTCTCTGGT   | AAAATAGAAA  | TAAATTGAAC  | 5700 |
| 35 | CGACAATCCA  | GTAATATAGAC | ATAAATACAG  | GATAAAAAAG   | CAAAAAGTTA  | AAAAATTGCA  | 5760 |
|    | ATTCTTTTAC  | CTACCTTTTG  | TTAGTTAGGT  | TGTAAGCCAT   | ATGGTAATTG  | ATAGTATTTT  | 5820 |
|    | AATTTGCAAT  | AGATTGTTGT  | TATAATTAAA  | CGGAAATATT   | TGTAATTGCA  | ACTTAATTTT  | 5880 |
| 40 | CCTGTAACAT  | AGTGTGATTA  | ATTTTCAGTA  | GGGGGTATTA   | AAAATTGAAG  | GATAAGATTA  | 5940 |
|    | TTGATAACGC  | AATAACCTTA  | TTTTTCAGTA  | AGGGGTATGA   | CGGTACAACA  | CTTGATGATA  | 6000 |
|    | TAGCTAAAAAG | TGTAATATATA | AAGAAAGCGA  | GTTTATATTA   | CCATTTTGAC  | TCGAAAAAAA  | 6060 |
| 45 | GTATTTACGA  | ACAAAGTGT   | AAATGTTGTT  | TTGATTACCT   | TAATAATATT  | ATTATGATGA  | 6120 |
|    | ATCAAAATAA  | ATCGAACTAT  | TCAATTGATG  | CTTTATATCA   | ATTCCTATTT  | GAGTTTATTT  | 6180 |
| 50 | TCGACATCGA  | AGAAAGGTAT  | ATTAGAATGT  | ACGTTCAATT   | ATCTAATACG  | CCTGAGGAAT  | 6240 |
|    | TTTCTGGAAA  | TATTTACGGA  | CAAATACAAG  | ATTTAAATCA   | ATCATTAAAGT | AAAGAGATAG  | 6300 |



EP 0 786 519 A2

TGCTGTTTCT TGAAGTGG TATTGAAAG CATCTTTTC GCAAAAATTT GGAGCAGTGG 6420  
 AAGAAAGTAA AAGTCAATTC AAAGATGAAG TGTATTCGCT ACTAAATATA TTTTGAAGA 6480  
 5 AATAATTTT GTTACTAGTT TGTAATAATT AACCTACTTT TGTAAAAAAA GACATGAGAT 6540  
 TATTTTITTA AATCTATATA AAGTTGACAA TACAAATCGA TATTGAGAAT ATTAAGATGT 6600  
 10 ATATGAATTT TATAAATTA ATGCAATACA TTAATATAAA TATCAATTGT TGCAAAATAC 6660  
 GATTGTTCAT ATGATTGAT AATATTATTC TTTATATTG TGAATGGTTA AGTTTGTCTT 6720  
 TGAACATAAT ATAAAAGTGT AATGTTCCC TGAAAAGAAT AAGTTGTCAT CTAATTACAG 6780  
 15 GAAATCCGCA TAAATTAGAT GAAATGGAAA GTAATAAGTA ATAATTTATT GATAAGCGCC 6840  
 TATGTGATGG TAAATCATGA CATAGGCGCT TTTTITTATA AGTTAAJAAT GTAAATAAAA 6900  
 ATTATATAAA TTACCCACAT CTTTTTAAAA GGTGTGGGCT TTATTATCAT TAACCAACT 6960  
 20 CACAGTGACG GGTATCGCAA GGTATTGAAT TACCGAGTAC GGGCAGCTC GGTGTGTAA 7020  
 AGAGCAATA ATCAAGTAAT GATGATGCTT CTAATCGATT ATAAGAAAGC CATGATAGAG 7080  
 TACGATGGTA TCTAGTTTTA TTAATTAATG GTTTGATAT TTAAGTTGG ACAATATTAT 7140  
 25 ATCTTGTGCA AAAATATAAA TAAGTTATAC ATAATGSTAT AGAATCATGA TATAATTTTA 7200  
 AACGATAAAA TATTATATA AATAATTAGA GAAAAATGTAG TTGTGTATGT yTTGTGGtCG 7260  
 30 TTAAACTAGA TATAATTGTC CGATTATATA AACATACATA ATGAATACaa TGATTGATTA 7320  
 TGTGGAGGAA ACCATGAmAG AmAAGTTTga TTTAGTAAAA CTATTAAATA TTCTAAAGAA 7380  
 GAATATTAATA TTATTGCTTA TTTTACCGCG AATATGTCTT GTAGTAAGTG 7430

(2) INFORMATION FOR SEQ ID NO: 261:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 4082 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 261:

ATGTGTTACT ATTATTTTT CATATTCACA CAATGATCT TGTTTATATT TAGCTAATTG 60  
 ATTTTATCT AGCATTTTAT CCTCTGCTG AGTTTGTAACT CTTTAATAAT TTTTCTTA 120  
 50 TAAAACTTA GTATTCCAGT TGCTTATTAT ATCATTGATG AAAGGCTGAA ATAAACATA 180  
 AACTGTTCCG ACCATTAACG CTGTAGCTAA AGATAAGTCT ACAAGTCCAC CTGTTTAA 240  
 TTGAATCGGT GTCTTCACAT TAAACGATAA gGaknAAAAAT AATTTCACGC CTTTGGTGT 300

|    |   |      |
|----|---|------|
|    | AGGCGTCTGA ATAATTGCA ATAAAAATGC TATGATTGCG ATAAATAATA TTGAATGCGT  | 420  |
|    | AAAGGTTGCG TGTCCAAAGA TCAATCTCAC AAAAAAACTA ATTACCTTAA ACCTTCTGCC | 480  |
| 5  | AATCTTACTT TGAGTGTGAC ATATATCGGG TAATAAGCTA GCTAGAGTTG CTAGAATGAT | 540  |
|    | AACCGTAACC GAGCAAAAAA TATCGGTTTG AAAATATTGT GTTGTTAGCG CTCCAACGAG | 600  |
| 10 | CATGCCGCAT GAAGCATGTG TTTTACCTGT CATATTGGTT CTCCTTTAAT ACTCACATTT | 660  |
|    | TACCACATCC CTAACAAAAA CAGCAACATA TTTTCGGGTT AAAATTCATT AGTATGACAC | 720  |
|    | AATTTAAAAA AGTATCACAT AACTCTTGAA AAGCAATACA AAATCGITTA TGATGTATTT | 780  |
| 15 | ACAAAATATT TAAAGGATGT GTTTGAATAA TGGCAATGAC AGTAAAAAAG GATAATAATG | 840  |
|    | AAGTGCCTAT TCAATGGAGA GTTGCTGATA TCAAAATTC TACAAGTGAA ATTAAAAATA  | 900  |
|    | TTACACAAGA CCAAGATATT CATGCAGTTC CTAATTTAGA CAGCAAAGAT GTATCTAGAA | 960  |
| 20 | TCGGCTCAAC GTTTGGTAAA ACGAATCGCG TTATTATCGA TACTGAAGAC CACGAATACA | 1020 |
|    | TTATTTATAC TCAAAATGAT CAAAAGGTTT ACAATGAATT AACTAAATAA ATTGTATAAA | 1080 |
|    | AAATCATTC ATGGTGAAGG CTTCATGAAT GATTTTITTA ATTGATTCAA CACCCAGCAT  | 1140 |
| 25 | AAACAAATAC AAAAGGACAA CTGTTCCCAT AATTTTAACA GTTGTCCTTT TTACATATA  | 1200 |
|    | TTTATAACAA AAGATGTGCC ATCAAAGAAA TAATTGGTAG TGAATGATT GTTCTAATCA  | 1260 |
| 30 | AGAAAATCAT AAACAATTTG CCGATGCTTA CAGGAATCTT CGAACCAAGT ATGACGCCAC | 1320 |
|    | CTACTTCAGA CAAGTATATT AACTGCGATA TACTAAGTGC CCCAATAACA AAACGAGTTA | 1380 |
|    | TATCATTTTG TACACCTTCA ATTAATATAG AAGGTAAAAA CATATCGGCA AAACCGATAA | 1440 |
| 35 | TAATGGTTTG AGAAGCCTGT GCCGCTTCAG GTATTTGCAT TAACCTTAAA AATGGAACAA | 1500 |
|    | AAGGTTTACC CAATATGACA AAAAAGGGCG TGTAGTTCGC AATAATGGTA GCAATAGTAC | 1560 |
|    | CAATACTCAT TACTACAGGC AAAATAACAA ACCACATATC AATGACTGTT TTTAATCCTG | 1620 |
| 40 | ACTTAAAAAA GTCAATAACG CCGGTGCTT TAATACCTAC TTCTGTTGCA GTATCAAAGC  | 1680 |
|    | CATGTCTCAA TGCCGTCTTT CTTCTGGCA ATGCTCAGT ACGCGCACTT TCAGGTACCT   | 1740 |
|    | CCTTAGCATA CTCATCAGGA ATTTTATTTA AAGGCCAAAT TCTTGGCATA ATGACTGCTG | 1800 |
| 45 | CAACGAGGCA GGATACTATC ACTGATAAAT AGAAAGCAAA AAATTGATTT TGCATGTGCA | 1860 |
|    | CTGTTTCAGC AACTACAATT GCAAGGTGA TAGAACTAC ACTAAATGTC GTTGAAATAA   | 1920 |
| 50 | CTGTTGCCCT ACGACGAGAA TAATATCCTT CACCATATTG TCTACTTGTA ATTAAGACAC | 1980 |
|    | CAACAGTTCG GTCTCCAATA AATGATGCTA AATTATCTAC CGTCGAACGT CCTGGCAATG | 2040 |
|    | TAAATAAAGG TCTCATAACC GGTCTAAAAA TAGGACCCAA CATCTCTAAC AAACCGTATT | 2100 |

|    |  |      |
|----|--|------|
|    | AACTTGAGAA CAATAATCCA CCCGTTTCAT CTGAGTAAAT AACCTTTGAA CCAATTCGTA  | 2220 |
|    | AAAAATGTCAT CCATGCAAAA ACAACTGCTA ATATTGCTAA AATTAACCAA CCAATTCTAA | 2280 |
| 5  | CGTTAAAAGC ATTGTTTCATT AGCCCGTCAG GTTTCATTT ATCTTTTAAA ATAGTTGAAC  | 2340 |
|    | AAATCAGAGT TATGATACCC GATAAAGTAA TTATCGTCAC AATTAATAAT GGCATTACGC  | 2400 |
| 10 | CACCTAATAC ATCTTTAAGC ACGCCTGCTA AAAATGCCAC GGGCAACGTT GTTTGCTTCT  | 2460 |
|    | GTCCATCTTG TTGCACTGGA ATTGGTACTA AAAATAATAA GATACCAATT AAAGACATCG  | 2520 |
|    | TAATAAACTT AAGTCTCCCA ATAACATCT CTTCCTTGA AAAGCTATCC ATAAATCAA     | 2580 |
| 15 | TCCATTTCTC TATGATTCG TTTTAAGTAT ATACAGAATT CTATTCAAGT AACAAACATA   | 2640 |
|    | TTCCCTATCA TTCTATCTTT CAAAATGTTT ATGTATGCAA AATAATGAAT AATTACAGTT  | 2700 |
|    | ATTAAATATA CGCTATTTCT TGTAATTTTT CAAGATGAAT TCAAAAAAGG TTAAGTACAA  | 2760 |
| 20 | TTACTGATTT CGTACTTAACT CTTTTTTAAA CTCTAATCAT ATGTTAGTGA TTTCATTCTT | 2820 |
|    | CGTAATAATA TTAAGAAGTA TGGTGCACCG ATAATTGCAA TGATAACCCC AACAGGAATA  | 2880 |
|    | TCCAGTGGCG GATGAATGCC ACGGGCTAAA CCATCTCCAA ATGTTAACAA TATAGCACCA  | 2940 |
| 25 | ATTAAACCCG ACATGATAAT AACGTGTAAT GTTTTATTT CTATTAAATG TCTCGCAATA   | 3000 |
|    | TGAGGTGCAA TTAATCCTAA AAAGCTAATA CCACCGACAA CTGAAATTGC GGATCTGCT   | 3060 |
| 30 | AATATTACTG CTAAAAATTAA CAATAGCATT TTAATAGTTT TAACTTTAA ACCGAGTGCG  | 3120 |
|    | GTGCAACAG CATCACCTAG ATTCAATACA TCTAATTGAT AACTCAATAA AATGATGATT   | 3180 |
|    | GGTATCGTTA TTA AAAACCA AGGTAATATA GTATAAATAT TCGACATATC ATGTCATAT  | 3240 |
| 35 | AGACTACCTG TCAACCAJAC AAGCGCTTTG TTTGCTTCCA GTGGATTCTT GATTAATAAG  | 3300 |
|    | AACTGCACAA TCGCCGTACA TATTGCGCCT ATTGCTAAAC CAATTAAGGC AAGCTTTGAA  | 3360 |
|    | CCTTTAACAT CATATTTTGA AATTAAAAAT GATAAAAAATA AACTTACTGC AAAGGCACCT | 3420 |
| 40 | AAGAAATGAAC CTATAGGTAA TACAAACAAT GGTGCTGTTG GAAAGGTCAT AATAATAATC | 3480 |
|    | ACAGCAGCTA AACTGGCACC TTTAGAAATA CCTATAACAT CAGGTGAGGC TAACGGGTTT  | 3540 |
|    | CTTATTACAG CTTGTATAAT TGCACCTGAA ATAGCCAAGC TACTACCGAT AATAATACCA  | 3600 |
| 45 | AGTAATGTTT TAGGTATACG ATACTCATTT AAAATAAAAT CATCTTGTTG AAAGATTCCC  | 3660 |
|    | TTAATAGCAT CAATCGGATG AATCATGACA GACCCTACAC ATAAACTTAT GAATATACTC  | 3720 |
| 50 | ACAATTAATA GGATGTGAT TAAACTATAA CGACGTATAA TTTTCGTTGT CATCATATTC   | 3780 |
|    | TTTTCACCCC TTTAATCGTT ATAAATAAGA AGTAAAGTGC ACCTACGAAT GATGTAAACA  | 3840 |
| 55 | TCCTACTCGT TGATTATATA GGATATGTAA TTAACGACT TAATACATCT GATAGTAGTA   | 3900 |

AGCGTTTGAC TATATGCGGT ACGATTAAGC CAACAAATCC AATTGGTCCT GCCACTGACA 4020  
 CCGACATACC TGTAAGAATA ATGACTAATA GTCCAATGAT AATCTAACT TTAITTTATAT 4080  
 TT 4082

(2) INFORMATION FOR SEQ ID NO: 262:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1145 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 262:

TATTAGAAGG TCGTTCGGAT GAACAATTAA AAAATTTAGT TAGCGAAGTA ACTGACGCCG 60  
 TAGAAAAAAC AACGGGGGCA AATAGACAAG CAATTCACGT TGTATAGAA GAAATGAAAC 120  
 CAAACCATTA TGGTGTGGCT GCGTAAGAA AGTCAGATCA ATAATCTTC ATAAGATGCA 180  
 TGCCAATTAA TTCTTTGAAA ACGAACAAGG CGACTICTAT CTGAGTATGA TAGAAATCGC 240  
 CTGTGTTTAT TTTAATCTTC ATCTAAAAG TCTTTAATAG CTTGTTTATT TGTGTTTTTA 300  
 TTAATCTGTA ATGCACTACC ATCAGTATTT GTATTGACAT CTTGTTATGA GTTCTTGATT 360  
 GGCACAGTCA ATGACTTAAC ATCTTTTTCA CCTCGGATAC CAAAACCTCAA ACCTGTTTGG 420  
 AAAATCCCTG AATCAGGAAT GTTTGTATTC ACATAGCCCT TTAATAATACC TGCAACTTTT 480  
 GGTAAATTTAA CAACGTGTTCT AAAATTAACC ATTTCTTTTT TCAATGTTTG CATCACTTGT 540  
 TCGTGAGCTC GCACGCGTCC GAAGTCACCT TCAGGGTCGT GACGGAATCT TGCATAACCA 600  
 AGTAATCTCT TACCATTCAA CCTATGGTTA CCCTTTTTCA AAGATACACC AATATTTTTT 660  
 GACATATCTT TTTCGACATT AATTGGTACA CCTTCAGGCA TTAATTCATC AATCATTTTT 720  
 TCAAAATCCAG TAAATCAAC TACTGCATAA TATTCAAGAT TAATTCCTAA ATTTTTATCA 780  
 AGTGTTTTTC TAAAGTACTC TGGACCACTT AAAGCGTATG CTGAATTAAT TTTGTGTTTT 840  
 CCATATCCTG GAATATCTGC ATAAATATCA CGCATGACAG ACATCATTTT CATCTTTTTA 900  
 TTGATAAAGT CATATTGAAC AACCATGATA GAATCTGTTT TTGATTGTCC ACCTGTGCT 960  
 TTATCTGCAC CGAGTACAAG AATAGAAATT ATACCATCAT TTTTACTGCG TCCATTAAAT 1020  
 TGATGTACTT TAACATCTTT CGCATGTTTC TTGGCATATT CTACACGCGT ATTGTAACCTA 1080  
 TGTAATATAT ATACAACATA TGCCGATAAG TAAATATACA ACAATCAGAA GAATGATAGG 1140  
 TAATT 1145

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 7075 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 263:

|    |   |      |
|----|---|------|
| 10 | TATGGCTCAT CATTAAIGCA CGTATCGGGT AGCGTTTACC ATIGATAAGT GCTTCATGTT | 60   |
|    | TAGCAGCAGT TCTTAAAAAT CCATCGCCAT AACCGATATC AACTACAGCT AATTTTGTAT | 120  |
| 15 | TGTTTTAGT CACTTCAAAG GCAAAGCTAT AACCGCAATA ATCACCAGCT TGTACTTCGC  | 180  |
|    | GCACCTGAAT AACATGTGCT TTTAAAGTTA ATGACTGAAC TATATCATGT TGATTCAGTG | 240  |
|    | AACATATAGG TCTTGAACCG TATAACGCAA TACCTACACG CGCATGTGTA TGGTGGGGTA | 300  |
| 20 | GTAATAYyyyg TCCTTCCCGA TAAAAATCG CACTATTTTG AGCATGGATT AGGTCGAACT | 360  |
|    | GATAACCTTC AGATAAAAGT GCCTCAACAA TTTCATCCA TTGTGAACGT TCAACATTAT  | 420  |
|    | AACTGACAC ATOGAATTCA TCAGCATATC CAAAATGGGk CCATAAACCA CTAATAATCA  | 480  |
| 25 | TTTTGCAAT TTGATTATGA TGGTGATCTT TCAATACTTC TTTAATTTTG TTTAAATCTT  | 540  |
|    | TAAATCCAGA CCGATGTAAT AAAATTTCAA ATTCTAAGTG AACATGAATA CCAGCTAAAT | 600  |
|    | CATTTTTATG GTTATAGTAA TATGTCAACG ACGSCAAAGT CATGTGTATT TGATGTTTAC | 660  |
| 30 | GGACTAAATC AAACTCGTAA ACTGCATTCA TTAAAAAGAT TGTGCTATCT GGAGCAAGTT | 720  |
|    | GTCTAATTGG AATTGCTTCT CGTAGTGATG TTGTGCTAAA TGTATCTATA CCTGCATGGA | 780  |
|    | TAAACTGAGT TACAGCAAAT TCTAGGTCAT AGTGATATGC ATTAATTTTA ACAACTGCCA | 840  |
| 35 | TTAATGGCTG ATGTGTTTTG ACTGTGATTG CATTTTGTA AATAATTTTC TTATTTCAG   | 900  |
|    | ACCAATGTTG TGTCATGTA TTACACCTCT TTGTAATTAT TTAATAAAAT TTCGTAACCA  | 960  |
| 40 | TTAACCACGT TTATTAACAC TTTTTCATCA AAATTTAAAT GTGATGTGTG CAAACCAGTT | 1020 |
|    | ACAAAACCTT TATCTTCATT TCGTGTTCCT ATAAAAACAA AGTAAGCTGG AGCTAGTTGT | 1080 |
|    | TGACCATAAA AACTAAAATC TTCCCCAAAT AAGAATGGCG TTGGTTTGTG ATAGACATTT | 1140 |
| 45 | AAATCAGCTT TTATTAAGGC GTCTCTTAT TTAGTACGTA ATTTCCGACT ATTGATTGTA  | 1200 |
|    | GGGGGATAAC CTTCGTGAAA TTAACTTCA CAATCTACAT TAAACAGAAG CTTGACACTT  | 1260 |
|    | TCTGCTATCT TGTCATTTG ATTTTTTAACG ATTGTTAAAT CATCAATATC ATATGTACGA | 1320 |
| 50 | ATAGTACCTT CTAATAGGCC ATTACTTGGT ACAGTGTTAA TCGCTTCACC AGCTTTAAAA | 1380 |
|    | TGACCAATAT GAACAATATT TCGTTTCAAA CCGTTAAGGT GAAATTGTTG AATTTGTGAT | 1440 |

EP 0 786 519 A2

|    |  |      |
|----|--|------|
|    | ACATGACTTG ACAGGCGTGT TAAGAAAAAG CGATACTCTG TTGCGCTGGC CGTAATTTCT  | 1560 |
| 5  | TCATCTCTTA TCAGTCAAT GCCITTCATCA GCAAATGGGT TAACATGAAT ACCAAATACC  | 1620 |
|    | GCTTCAATTG GATACCTTATC AAAGGCACCG GCTTTTATTA ATCGATTGTC ACGGCCACCA | 1680 |
|    | GTTTCTTCTG CAGGTTGGAA AATGAAAACG ACATTTTGGC GTAATTGACC TGCATCTTGC  | 1740 |
| 10 | ATGTCTTTGC AACGTTGTAC AAAAAGCAAT AATGCAGTTG TATGACCATC ATGTCCACAA  | 1800 |
|    | GCATGCATCA CATGATCAGA TTGACTGCGA TAAGGCACAT CATTTCCTC TAAATAGGT    | 1860 |
|    | AACGCATCAA TATCAGCTCT ATACGCTATC GTATGTGAGC CATTACCTTC TAAGTATGCA  | 1920 |
| 15 | ATGACGCCAG TTTCCAATGG GCAATCGTAT TTAATATTTA AACTATCTAA AAACGCTTTA  | 1980 |
|    | ATATAAGCAG TTGTTTCAAA TTCATGTAAG CTTAATTCAG GATGTTGATG TAAATGACGG  | 2040 |
|    | CGATGTTTGG TAACAAATTC TAATTCATTG ATAATTATCA ATCCCTTTGTG TTAATTAAT  | 2100 |
| 20 | ATATAAATAG TGTAACTGAT TTCGAAATTT GTGATCATAA GTTTATTCAA TGCTAAACAA  | 2160 |
|    | TAAGGTTGAG ACATAATCGT ATCTCAACCT TGAATATTAT ATACGTTGAC GTCAGTAGTC  | 2220 |
|    | ATTGAGTTTT CTTAATGCTG CTACAATCTC TTTTTAGTA TCTTGTAATT CAGAAGCTTG   | 2280 |
| 25 | CTTAATCACT TTTGCAGGTG TACCAGCAAC AACTGCACCA GCTGGTACAT CTTGTGTCAC  | 2340 |
|    | AATCGGCCCA GCTGCAACAA TAGCACCTTT ACCAACAGT ACACCTTCTA AAATAACTGC   | 2400 |
| 30 | ATTTGACCG ATTAATACAT CATCTCGAT TATAACCGGT GAAGCACTAG GGGTTCAAT     | 2460 |
|    | CACACCTGCT AATAGTGGC CAGCCCTAC ATGTACATTT TTACAGTTG TAGCAGGACC     | 2520 |
|    | AOCGAGAGTA GCATTATAT CAATCATTGT ACCTTGGCCA ACGACTGGCG CAATATTAA    | 2580 |
| 35 | TGTTGGCCCC ATCATAACGA CAGCACCATC TTCAATAATG GCTTGTCTC TAATAAAGC    | 2640 |
|    | ACCTGTTTCA ATTGTCGCAT TCGTATTGT TAAGTCTTTT AATGGAATAG CAGAATTGCG   | 2700 |
|    | ACGAACCAAT TCAATTTCTA TATCTCGAA TTGACTACCA TATGCTTCTG AAAAAGGTTT   | 2760 |
| 40 | CCAATCATCC GCTTCACAAA AGATTACTTT AGATTGTTCT GAACCAATA CTTTAAACT    | 2820 |
|    | TCTGGATAT GTGATGCCCT CAAAATTACC ATTAAATAT ACTTTTATTG GTGTAGACTT    | 2880 |
|    | TTTAGCATCA CTTATATATT GAATAATTC TTCAGCTGTT AAATGTTGTA CCATAAAATA   | 2940 |
| 45 | ATGATCTCC TTTAATATGT TTATAAGTTG TCAACGTAT AAAAGCCGTT TGGTTTATTA    | 3000 |
|    | ACTAAGCGTT CTGCTGCTTG TATTGCACCA TTGCAAAAA TATCTTTTGA TTGTGCAGGA   | 3060 |
| 50 | TGCGTAGATT GAATCGTTTC ATCAGTGCCA GCAAATAGAA CTTATGTTTC ACCGCAATC   | 3120 |
|    | GTACCTCCAC GAATAGAATG TATACCAATA TCTGTGGCT GCGGTTTTTC ATTTAATTCA   | 3180 |
| 55 | TGCTATCAT ACACAGGTGT TACATTTCT TTCAAAGATA CGATCACATC ATACATTTT     | 3240 |

|    |            |            |            |            |             |             |      |
|----|------------|------------|------------|------------|-------------|-------------|------|
|    | TCGAATCAT  | CAAGTAGGG  | AACAGCAGCT | GCTAAAATTT | TAGTCAATGC  | ATGAACGCCA  | 3360 |
|    | TAACTCATGT | TCGCGCTGAA | AAACACAGGC | ATATTTTGAC | TCAATTCAATC | TAACATTATTA | 3420 |
| 5  | AGTAGTTTTT | CTTTCTCGnC | CAGTgTTGCC | ACAACTAATG | GCAAAATGAAA | ATCTTCATCT  | 3480 |
|    | AATAAAGGGA | AAAGCAGATT | TGGATTGAA  | AAATCTATTG | CAACATCGGC  | ACCTTTAACA  | 3540 |
| 10 | CTGCAATAT  | GTGATATTG  | TTGATATGCG | GTTGTGCTT  | TCGGTGATT   | TTCAATGACC  | 3600 |
|    | CCAAGATT   | CATGTCTTT  | TTCTTCTGCT | AATCTAGCAA | CGCGTTGATT  | CATTGCGCCA  | 3660 |
|    | TAGCCAATTA | GTAATATTTT | CACTCATTTT | CACCGGCTTT | AAATGCTCA   | TATGTTTCAC  | 3720 |
| 15 | GAAGCACTTT | AGTATCTGTA | TCTTCTAGGC | TAACCAATGG | TAGACGTAAT  | TCATAATTTT  | 3780 |
|    | CAAATCCTAA | ATAACTTGTT | AGAGCTTTAA | TAGGAATTGG | GTTAATATCA  | ACTGATAAAG  | 3840 |
|    | CTGATAACAG | TGTGCGGATT | GGTTTAAATT | GATCTTGAAT | ATCTAATCCA  | CTTTGTTGAG  | 3900 |
| 20 | CATCGTATAA | CGCTTGAAT  | TCTTTAGGAA | TGACATTGGC | AATAACAGAG  | ATAACCCCTT  | 3960 |
|    | GACCGCCACG | TTGATAGTAT | TCGACGACGT | TGTCATCATT | GCCACTATAT  | AATGCAAAATG | 4020 |
|    | AATTTGTATC | AATGCGCTTT | TTCACTTCTT | CTAAATACTC | AAAATCATTG  | GTAGCATCTT  | 4080 |
| 25 | TTAAAGCAAC | TATATAAGGA | TGTTGACTTA | ATATTTCTAC | AGTTTCTGGT  | TCAATTGTCA  | 4140 |
|    | TGTTCGTTCT | TGAAGGAACA | TTGTACAGCA | CGACTGGTAA | TTTCACAGCA  | TCTGCAATCG  | 4200 |
| 30 | CTTCAAAGTG | TTTGACTAAA | CCACGTTGGT | TCGTTTGT   | GTAGTAGGGC  | GTAATTAACA  | 4260 |
|    | TAATTGCATC | AGCCCCTAAG | GCTTTAGCTT | GGATTGAAGC | TTGGATTGAC  | TTTTCAAGTAT | 4320 |
|    | CATTAGTGCC | AGTTCCTGCT | ATGACAGGAA | CACGTTTATC | TACAAGATCA  | ATAACTGTTT  | 4380 |
| 35 | TTAGAATGCG | TTCTTTTTC  | TCTGTTGTTA | AAGTAGGGCT | CTCAGCAGTA  | GTCCATTAA   | 4440 |
|    | CGATGATTGC | TGCGGCATTA | TTTTCTAGTA | AAAAATTAAC | GTGTGTTTTC  | AAAGCTTCAA  | 4500 |
|    | TATTTACTTT | GTTATTTGTA | AAAGGGTTG  | TAAGTGCAAC | sCCAACACCC  | TCAAATAAAT  | 4560 |
| 40 | GTGTCTTTTT | AATTGCGTCC | TTTTAAACGC | ATAACTTGTT | CCAATACTTG  | TACAGCAITTT | 4620 |
|    | AATGCAGCAC | CTTTTAATAA | ATTGTCTGAT | GTACACCATA | CATGGAAAAGT | ATTTCTTAAT  | 4680 |
|    | GAATCATCTC | TACGTATACG | GCCAACAAAC | ACTTCATCTT | TATTAGTAGA  | ATTGATTGCC  | 4740 |
| 45 | ATTGGATATT | CATTGTTCTC | TGGATTGTCT | ACTAAAACAA | CGCGGTCAATC | TTGATCAAAT  | 4800 |
|    | AACGCTTTAA | TATCTTCTGC | TGTTGTTTCT | TTGTCAAGCG | TTACATCAAT  | TTCAACACTA  | 4860 |
| 50 | TGACTATCTT | GAACAGGCAC | ACGTGCGCAT | GTTGCTGTTA | CTTTTAAATC  | TGCGCATTTT  | 4920 |
|    | AAAATTTTTT | TGCTCTCATC | AATCAITTTT | TGTTCTTCTT | TTGTATATCC  | GTTTCTTAAA  | 4980 |
| 55 | AACACATCAA | TATGCGGTAA | CACATTATTA | TAAATTGGAT | GTGSAATGTC  | TTCTGTGTCT  | 5040 |

EP 0 786 519 A2

|    |  |      |
|----|--|------|
|    | TGATATGTTG TATATGCCAC TCGTTTTAAA CCATAAGCAT CTGCAATAC TTTTAGAGGT     | 5160 |
| 5  | ACAACAGATT GAATCGTAGA GCAGTTTGGA TTGCAATGA TACCTCTGT AAATGTAGGT      | 5220 |
|    | TCATTGACTT CGGAACGAT TAAATCAATA TCTTCTGCCA TACGCCATG ACTTGAATTG      | 5280 |
|    | TCTATAACGA TTGCACCAGC TTTTTCAAAA AGTGGGGCAA AGTGTTCGCT TGTACGCCCA    | 5340 |
| 10 | CCAGCACTCA TTAATACATA ATCGAAATGT TCACTTGCAC GAGCATCAGT TAATTCTTGA    | 5400 |
|    | ACTGTATATG TTTTTCCTTG AAATTCACCT TCTTGCCTG CAGAAGGTGC TGATGAAAAT     | 5460 |
|    | AATACTAATT CATCGAAAGG AATATTTTTA CGATTAAATG TCTCCAACAT TTTTGTACCT    | 5520 |
| 15 | ACTAATCTGT TTGCACCCAC AACTGCTAAC TTTGTCTATA CTGTCTACTC CATTTTATAA    | 5580 |
|    | TAATTTCcAA TTTTITAGAAT ATTITTAACAA TCATTTTACC ATTTAAATGTT AAATGCGTCA | 5640 |
|    | TATAGTTTTT CTACCGCTTG TTGCCCATTA AAATCATCAA TGACGTATGA AATACTTATT    | 5700 |
| 20 | TCAGATGTTG TTGTTTGGTA GAAAGGTATA TTAATTTCAA TTAATGTCAA AAATGCTTTT    | 5760 |
|    | GATGCCACAC CTGACATATC ACGCATGCCT GAGCCAATTA ATGAAATTTT GACATAATGC    | 5820 |
|    | TCATTGATTT TATAAGCTAA TGCTTCATAT TGATTCTTTA ATGTTTCAAG AATCATAGAA    | 5880 |
| 25 | ATTGTGATGA AATCACTATC TTTAATCGTG AAGGATAGTT GTAGCCCATC CAAGTGTAGC    | 5940 |
|    | ATTGTGTGAA TCATATCAAC ATTTACAGCA CCTTCTTCAA GTTCCGTAAA TAGTTGGGTA    | 6000 |
| 30 | AGTAGCTGAT TGTACGGTAG GGGATAACTA ATTGTTACAT GCATCATATG TTTATCCAAA    | 6060 |
|    | GCCACACCAG TAACTGCTTT TTTCTCTAAT ATTTCTTCAT TTGACATAAT CCATGTTCTT    | 6120 |
|    | TTACGTTTCG ATAAAGTTTT TCCTAAATAT AAAGGGATAT TATAGTTTTT AGCTAATTCA    | 6180 |
| 35 | ACACTTCTTG TTTCAAGTAC ACCAGCACCT AAAGCGCTCA TTTCCATCAT TTCTTCATAT    | 6240 |
|    | GAGACAGATG CTAGTCGTTT AGCCTTTGGT AAAAGTCTTG GGTCAAGTGC ATACACACCA    | 6300 |
|    | TCATGCTCGG TATAAATTTC ACAAGGTATT TGATTACTAA CAGCAAGTGC CACAGCGGTC    | 6360 |
| 40 | GTATCAGAAC CACCTCTGCC TAAAGTTGTT AATTCTGAT GTTCATTGAT GCCTTGAAAT     | 6420 |
|    | CCAGCAACTA CTAAAATATC GTTTTCTTGA AAGGCTTGTT CAAATGTTTG AGGATTAATT    | 6480 |
|    | TGAGCAATTT TACTTTTTAA ATGATGGCCA ATGGTTTTAA TACCCGCTTG ATAGCCAGTC    | 6540 |
| 45 | ATTGCTTTGG CATTCAATCC GATATCAITT AATACCATG ATAAATAAGA TACAGTTTGT     | 6600 |
|    | TGCTCTCCGG TTGTCATAAA TAATGCCAGT TCTTGTGTTT TTGGTGTCTT AGTCAAGGTT    | 6660 |
| 50 | GATACATTCC TCATTAAATTG ATCTGTTGTG TTACCCATAG CACTTACAAC GACAATTAAa   | 6720 |
|    | TTGTTCTACT TGAATTGACT GCTCCTTTAA CATTTACAGC ATCCTTTTTA TTTTGTGIAAA   | 6780 |
| 55 | ATCACTGAGC GATGATCCGC CAAATTTCAA CACACTTCTT GTTACCATAT AATCCTCCTA    | 6840 |



# EP 0 786 519 A2

TAATCTATAT ACAAGTGATG CACTCCATTA TTTTAAATA ATGACAACT CTCAGCTCTT 6960  
AACC AAAAG TCCAACAAAT TATAACTGCT ATTATAATTG CTTCGGCATC GCACCCCTTC 7020  
AAATTTAGCT GTTAGCAGAC AGTAATCTAa ACTTIACTCA TGATTGATGC GCCTC 7075

(2) INFORMATION FOR SEQ ID NO: 264:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 5171 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 264:

AGACGTACTT TGTGATTmCG AAGyrCGTAC TmAGCACITG GTCGACGTTG ATGTACTTGT 60  
TGAACCTGAT TGACTAGTAC TTTGTGATAA TGACTTACTA TCAGAATCAG ATGTACTTTG 120  
TGAATCACTT AATGATTCTG ATGTACTACC TGACTGAGAC GTGCTCATTG AACTACTTAC 180  
GGACATTGAT TTACTGTCTG ATGCAGATAA TGACCCACTT GTACTGATAG AGTCACTTAC 240  
TATCTCTGAA GTACTCATCG AGTCTGATGT ACTTGTTGAG ACACTTTGTG ATGCTGCTAT 300  
GCTTAGTGAT CCAGAAACAG AACCCTTGT GCTCGTCGAA TCGCTCAATG ATTCTGATGT 360  
ACTCATCGAT TTGAATCAC TTGTACTTAA TGATATTGAT GTACTTTGTG AATCTGATTT 420  
GCTTGTTGAC GCACTTTGAG AGTTGGCTAT GCTATTIGAA ATACTGATAG AGTCCGAGGT 480  
GCTAGCTGAC TCGCTCAATG ATGTGATGT ACTAATIGCA TTCGATGTAC TGTCACTTAA 540  
TGATGCTGAT GTACTAGACG ACCCTGATAT ACTCGTGTAT AAGCTTTGTG ACTTAGACAA 600  
GCTTCTGAT GTACTCATAC TTAATGAGTC ACTGAGTAT GTTGATGTAC GCAATGAATC 660  
AGATGTACTT GTTGATAGAC TTTCGGATT TTCACTAGT CTAGAGTTTG AAATAGAATC 720  
GCTTAATGAT GTTGATTAC TAGCTGAATC CGACATGCTT GATGATACAC TTTGTGAATT 780  
CACTAAACTT GTGCTTGTG AGCTTGATAC ACTATTACTT TCAGATGTGC TTAATGACTT 840  
AGATGCACTC ACAGATCAG ATAGGCTTAC ACTTGTGAT TTCGAGGTAC TAGCTGATGT 900  
AGATACCACA ATCGATCCTG ATGTACTCGT TGATGCACIT TGTGAGTCAG CTTTACTTGT 960  
TGACACACTT TGAGATTGTT GTGTACTTCC TGATGTTGAT ACGGAATCAC TCATGCTATT 1020  
TCTTGTTACT TCATATTAA AAGTTGTGCT CGTTTGTGTA CGCTCGCAT CTGTAGAAAC 1080  
GATTGATATA GTACTTGTAC CAATGTTTGT TGGTGTACCA CTAATAGTAT TATTGTACT 1140  
ATCAAATGTT AGTCGGATG GCAATCCAGT CACTGTATTG GTCACCGCAT TTCCACTGTT 1200

|    |  |      |
|----|--|------|
|    | ATTGGTGICA CTGTTGGTGC TGTGCTATCC ACAACATTTA TTGTAAAGT TGTGTCGAT      | 1320 |
| 5  | TTGTTATTATG CTGGTGCAGT AGACACAACCT GTCACTGTTG ATTGACCAAT TTTTGTGTGGT | 1380 |
|    | GTCCCAATGA TTGAATTCGT TGCACATATCG TAACTTAATC CGCTTGGTAA TCCTGTAACCT  | 1440 |
|    | GTATTTGTCAC CAGTCCACGT ACCATTATCC GTTGTAAGTCA ATACAATAGG ATTCATTGTT  | 1500 |
| 10 | TTACCCACTT CTATGGTTTG ATTGCGCTACA GTTACAGTTG GTGCTTTTAC ATCAGTAAAA   | 1560 |
|    | TAATATGTCA CTGATTGTCC AGCATTCTGC ATTTTACAG TTTTATTGTT ATCATTATAA     | 1620 |
|    | GTTGACGCAT ATGAACATC GACGGACGTG TAGTTATATC CTTTAGCAGT CAATGCAGAT     | 1680 |
| 15 | TGCTGATTAT CGATTGTAC GACTTGATCA ACATTTCTGT AATATGTTTT TGGTGGAAATA    | 1740 |
|    | ATATCTTTAC CTGTTGTAC ATCAACGTAT CTCACTGTG TATCAGCAGA CTCTGTATAT      | 1800 |
|    | TCGAATGTTT CAAATTGTAC TTGTTGTAAA TTTGTGCGCG CACCTGTTGA GGTGTCATT     | 1860 |
| 20 | GATAATGAAA AGTTGGTCGT ACCACTTTTC GCAATCCAAT CTGAAATATT ACGTGCCAT     | 1920 |
|    | GTTTGACCTG CATATTGAC AGTCATAACC TTTGTATCAC CATTATAGTT AATATCAAAA     | 1980 |
|    | TCTTGGaACG TGTATTGTT AGGTTGaACA TTTAACTTCG CAGCATTATC AGCTGTTGAA     | 2040 |
| 25 | CTTGATGTAT ACGTTGTGCG AACACCATAA CTATCTGTTG TTACAAATGC ACCAAACGCA    | 2100 |
|    | CCTCCACCAG CTACATTAGA TGGGTGAGCA TTGCGCTTGT CAGCTGAATT TGGTTTAGAT    | 2160 |
| 30 | GTATTGTGAT ACGTATCCAA TTTGAAGCCA AATGCGTTAC TTAAGCCACC AATACCTACT    | 2220 |
|    | GCGGCACCGT TTAACCTGT TTCACCTAAT ACACCTGGTG AAAAGGCCAA ACCGATACCA     | 2280 |
|    | TCTCCACCAT TTCCATGCCC TTCATATTG TTACCTAAAT TTACTTTTCC AGAAAAATGA     | 2340 |
| 35 | AAACTCTTAT TAGAGTCAAT ACGTGTTCCT AATGTAATAG CACCTTTTGT GCTGTATGCA    | 2400 |
|    | TCCTGTGTTA ACGTCACAAT ACCGGTACTT TGATCATAGG TAGCATTACC TGACGTTGTC    | 2460 |
|    | ATATATTGTT TTAAGTTATC TTTATTAACT GTAATTGSTAT TAGCAGTTAC TGCGGTTGTC   | 2520 |
| 40 | GTGCTGCTG ACGCAAATGT TGACATAGCT AAGCGACTGA AAGTTCGAAG TTTTACTGGT     | 2580 |
|    | GCGGTGCTAG TTGACGTTGT GCTAGTTTTG TTTAAGTTGA CCGAAGATGG CGTTGTGCTT    | 2640 |
|    | TGTGAAGTGT TATTIGATGC AGTACTTTGA TTTGTTGATG TATTAAATGG TTGTTCTGTA    | 2700 |
| 45 | CTTGAAGTTG AAGCTACAGA TTTAGTATCA GAACCTGATG TAGTATTCTT TGAGGATGTT    | 2760 |
|    | GATTCTGATG TAGATGTCAA TTTCTCTTGT TGATTGCTTG TACTATTAGT TGTGGAAGTG    | 2820 |
| 50 | ACCTTTTTCAG ACTTTTCACT TGAGACTGTG TCACATTTTG ATGTTTGTAC CGAACTACTA   | 2880 |
|    | TTTTCGTTA CACTGTGGA ATCGGCTGTT GATGTTGATG CTTCGATTGT CGTTGAGTTT      | 2940 |
|    | TGATTACCTA CTGTTTCACT TTGTTGTTT AATTCAGAAG TTAATGGTGC ATCAGAAGCC     | 3000 |

|    |             |            |            |            |            |             |      |
|----|-------------|------------|------------|------------|------------|-------------|------|
|    | GTTCCTCAGTC | CGTATCCCGT | CATTTTTTTA | CTAATGCTTT | GATTATCTTG | ACTCACTAAA  | 3120 |
| 5  | CTATGACTAA  | TAAATGGTAG | CCCATAATT  | TGGAACATT  | CTATTTCTTT | AATTCGGGAT  | 3180 |
|    | TTTACCCAAT  | TTTTTCGAGA | TTTATAAAGT | CTTACTCTTG | TTTTTCGGT  | TGCTAAGCTG  | 3240 |
|    | TCATGAAATG  | CTTTCTGTCT | TTTACTCATG | TAATAACTCC | TTGTATTATC | TTTACATTCA  | 3300 |
| 10 | TTAGATTATA  | ATATATGCCA | CTATTCAATT | TAATACAAC  | CTTTTTTGAT | ACAAAAATAC  | 3360 |
|    | TCATTITGTT  | AAAATTGTGA | AAAATTCAAT | TTTATTCGTC | TAAATGTAAT | CGTTTTTCATA | 3420 |
|    | TTTTTAAAA   | TACTTTTTCT | CGTTTATGCG | TATAATCTTT | TTTTATATAA | ATTGGCTGTA  | 3480 |
| 15 | TTGGCTTTAT  | GTTTAATCAT | TATAATTGTT | TCGTTTTTAA | AATAATTATT | GTATTAAATAT | 3540 |
|    | ATCTATACCA  | TCCACCTTTT | ATTATATAAT | AGTTAATTTA | CAACTAAACG | ATAAAATATA  | 3600 |
|    | TATGCAAAAT  | ACATCTTTAA | TATTAAAGTA | ATACCAATAT | TTTTTCAATA | AACCTAGTGT  | 3660 |
| 20 | AATATATGTG  | TAATTCTAAA | AGATTCTTCT | TAAAAAATAT | AAATACCACG | ACATATTGCT  | 3720 |
|    | TTAACATTTT  | CATTTATAAA | GCGAAAAAAT | GCATCGCTAC | TAAGTTGAAT | GTTTAGTAGA  | 3780 |
|    | GATGCATTGA  | ATTCACATAA | ATGATTAAAT | TACTTATATC | TTTTCATCTG | ATTGATTATC  | 3840 |
| 25 | GAATTTCTCT  | CCTTCTAAAC | CTGCTAACTC | TTCTTTAGAA | GCTGCAGGTG | CTTTCATTTC  | 3900 |
|    | AAATATCTCA  | TTCACCTACT | TGTAATCGTA | ATATCCTAAT | CTGGCAATAG | GTTTAATCGA  | 3960 |
| 30 | CTTAATGTCC  | AATTTACCAT | TATCAAGAA  | AACCTTATCG | TCAATATGAA | CTTGGGCAAC  | 4020 |
|    | TCTTCTTATA  | ACAATATCTA | CGGTAGATAC | TGGATCTCCA | GTTGGAATAC | GAATCGTTTG  | 4080 |
|    | AACGTACTCA  | CATTCAAAAT | GAACGTGGCA | TTCTTTTACA | CGATATCCTG | GAGCTTCTAT  | 4140 |
| 35 | ACATTTTTCC  | TTTGTTACAC | CTGCAAAAT  | AAATTCATCC | TCTTCTGGTG | GCAATTGCTT  | 4200 |
|    | CGATGATAAA  | TTAACTGCTT | CTCTTAAATC | ATACGTTGCC | ATATTCACCA | CAAAACCAAC  | 4260 |
|    | TGTCTCTTCA  | GCATTTTTCA | CTGTATCTTT | AGTTGCTGTA | TCACCAAGAA | CGGATTGATT  | 4320 |
| 40 | TGCTGCGAAC  | ATAACCATAG | GCGGATCCCA | AGTTAAGTTT | TGATACTGAC | TATAAGGCGC  | 4380 |
|    | TAAATTATCT  | TTCCCATCTT | TCGATACAGT | AGAGATCCAC | CCTATTGGAC | GTGTACTGT   | 4440 |
|    | ACTACTTTTA  | AATGGGTGCT | GCGGTAAACC | ATGACTTCTT | ACACCTTGTT | TTGGCGAATA  | 4500 |
| 45 | ATTCATACTA  | TCTTCACCCC | TTATAAGTAA | TTACATTTAA | GGTACGCCCC | TCTTTACATA  | 4560 |
|    | AGCGTCTAAT  | ATAAATAAAC | AATTTATTTA | TAAGTAGAAA | CTATATATGA | CGTGGTGTCT  | 4620 |
| 50 | TATAATTTCG  | GTTCTTGATT | CGAAAAATTC | AGATAAGGAT | TTATACAATT | AATATTTATG  | 4680 |
|    | ATATCTTTTG  | TAAATTTAAT | TAATTATAGT | TACTTCAATC | ATGATTAGTT | TATAATAATA  | 4740 |
| 55 | AGTGAAATT   | GAATAAGACA | GCTATTATGC | GATGAGCGAA | AAACTTCAAG | TAAACAAGA   | 4800 |

TAAGTTCAAA AAAGAATTCA AACCTGTTAT GCACTTAAAA GGTGATGCAT TCAATCAACA 4920  
 GTTACAATCT TTGATTACA AATATCCACA AATACAAAAA AATATGAAAT CAGAGTTCAT 4980  
 TGCATTATT GATAAGAAA AAAATAGAGA AACAGTAAAA AACTATGCTT GGAACCTTCA 5040  
 AAAATCTATA AATGACATTA TGCAATCATA TCCTAGCACA AAATTTGTAC AGTTTITATAA 5100  
 AAGATGATGT TTCCCCGTCA ATGGTAGATG GAAATGGCCG TTTAAAAATCG GGATACTAAT 5160  
 GTATTTCCAT C 5171

## (2) INFORMATION FOR SEQ ID NO: 265:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 3589 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 265:

CTACACACTA AACCTATTTT AGTTATGGGT GGTACAGTGA TTCTCTTTTC ATTTTAAATA 60  
 GGTATTTGGA TTGGTCATCC TATTGAAACA GAAATCAAAC CACTTATTAT TGGTGCAGAT 120  
 ATTATGTACG TACTTGGGCT TGTAGATGAT ATCTACGATT TGAAACCGTA TATAAAATTG 180  
 GCTGGTCAAA TTGCCGCTGC CTTAGTAGTT GCTTTTATG GTGTGACTAT TGATTTTATT 240  
 TCGTTGCCAA TGGGTACAAC GATTCATTTT GGATTTCTTA GTATTCCAAT TACTGTGATT 300  
 TGGATTGTTG CTATTACAAA TGCAATTAAC TTAATTGATG GACTCGATGG TTTGGCGTCG 360  
 GGTGTTTCKG CAATCGGACT CATTACAATA GGGTTCAATT CAATTTTACA AGCTAATATT 420  
 TTCATAACGA TGATTTGTTG TGTTTTATTA GGCTCTTTAA TTGGGTTTTT ATTTTACAAT 480  
 TTCCATCTTG CCAAAATATT TTTAGGTGAT AGTGGGGCTT TAATGATTGG ATTTATCATC 540  
 GGATTCCTTT CTTTACTCGG ATTCAAAAAT ATTACAATTA TTGCATTGTT CTTCCCAATT 600  
 GTTATCTTAG CAGTTCCATT CATTGATACT TTGTTGCGAA TGATTGACG TGTGAAAAAA 660  
 GGGCAGCATA TAATGCAAGC TGATAAATCG CATTTCATC ATAACTATT AGCTTTAGGC 720  
 TACACACATA GACAAACAGT ATTATTAATC TATTCAATCT CTATTTTATT TAGTCTTTTCG 780  
 AGCATTATTT TGTATGTATC GCCACCATTA GGTGTTGTAT TAATGTTGT ATTAATCATA 840  
 TTTAGTATTG AATTAATTGT TGAATTTACA GGATTAATAG ATAACAATA CCGACCAATA 900  
 TTAATTTTAA TTAGTCGTAA GTCATCTCAT AAAGAGGAAT AGGGAATGAA AGCATAGCTG 960  
 ATGGGATAA TTTGTATTAT ATGGCTTTAC TCTTTACAAT TTTTGTGTAT TAAATTTCAA 1020

|    |  |      |
|----|--|------|
|    | ATTACCGTC TTATGATAGT GCTTTTATT TTTATTCAGT TGGTATATCG AAAGGTAAC     | 1140 |
|    | GCTTTGGAGT TTCTTCAGTC AAATCGAAAT TTCTGTCAGT CATTTGATT AAAAGTTAA    | 1200 |
| 5  | TAAACGGTTC ATAGTCACCT TTAACGACAT CGATATAGTA GCTTACCTTA TCAGTGTAAG  | 1260 |
|    | TTTGGTTTCT TAACATAAAA TGAGTTGAAG CTAATTCATA TTCAAATTTA CCAGTTTGAT  | 1320 |
| 10 | CATAATTCAG TGTACTATA CATGGTACTG CTTCTGCTAG TTCGACACGC CCGATATCAT   | 1380 |
|    | AAATGACCTT TCTAACAGCA CCGCTATAGG CGCGAATTAA ACCGCCACCA CCTAAATTTAA | 1440 |
|    | TACCACCAAA ATATCTTGTT ACTACGACAC ACGCAATTAG AACATCGrGC TTTTTTAAATA | 1500 |
| 15 | TGCTCAACAT TGGGACACCG GcAGTTCTcG TCGGTTCCAC ATCATCATTC GcNTTITGAA  | 1560 |
|    | TATTCATTcC AGGTCCAATA GTATATGCAG AACAAATTAT AGTGGCATCT TTATGTTCTT  | 1620 |
|    | TTTTTATTGC AGCAATAAAAT GCTTTaGCTT CATCTTCATT TTGAACAGGT TTGATATGAG | 1680 |
| 20 | CAATGAATCT TGATTTACTA ATCACATTTT CAATAATGTC TTCTTTTTTA ACAGTAATGA  | 1740 |
|    | TATTTTGIGT CATAATAACT CCTTAATTCA TAAGCTTAAG ATTATTTAAT CTTCATTATA  | 1800 |
|    | CACGTAAAAA GACATGACTA TAAATCGTTT GATTGCCATT TTCTTTTTAA CTGAAATATT  | 1860 |
| 25 | GTATCATTCG TATGAGTATA TTTTAGGAGG ACGACTATGA AAATTGCTGT GATGACCGAT  | 1920 |
|    | TCTACAACTT ATCTGTCGCA GGACTTAATC GATAAATATA ATATTCAAAT AGCGCCATTA  | 1980 |
|    | AGTGTGACTT TTGAAGATGG CAAGATTATA CCAGAAGAAA AAGTTCGTAC TAAAAAGCGT  | 2040 |
| 30 | GCCATTTCAA CATTAGAAAA GAAAGTATTA GATATTGTAA AAGACTTTGA AGAAGTAACT  | 2100 |
|    | TTATTTGTCA TAAATGGAGA TCATTTGCGA GATGGTCAAG CGTTATACAA AAAGTTACAA  | 2160 |
|    | GATGATTGTC CTTGAGCTTA TCAAGTAGCA TACTCTGAGT TTGGTCCAGT TGTGACGCA   | 2220 |
| 35 | CATTTAGGTT CTGGTGGATT AGGTTTAGGC TATGTTGGCA GAAAAAAG ATTAACATAA    | 2280 |
|    | TTATAAAAAT TTAATAAAAG AGTCTATATT GTAATTGGAA ATTATCTCTC GTATACATGG  | 2340 |
| 40 | CTTTAAATGT TCATCATTTG AAAGCCAAAA TGCTAAAGAT ATAAGAAAAA CATTATAATA  | 2400 |
|    | TTAGGCTCTT TTTTACGTTG AAATGAGGTT TTAAGCATTAA AACATTACGG GAAATTAATT | 2460 |
|    | CATCTCATA CTTCACTTAC TAATGAAAAA ATTAAAAAAG AAGTAACAGG TGTATCAAAA   | 2520 |
| 45 | CAAAATTCAA ACTATTATTG TGTCAATGT GAAAGTACAA ATCCAAAGCA TTTTATCAG    | 2580 |
|    | TATGATTCTT CAGTACATTC CAAGAAAAAT GTATATTGCA GAAATTGTAT ATCATGGGT   | 2640 |
|    | CGAATGGATA ATGTAAACAG ATATAAAATA ACAGAGAGTT CGCAAAGTTC ATCAACAGCA  | 2700 |
| 50 | TATTATCATC TCTCATTTGA ATTGTCGGAA CAGCACTCTT ATGCCCTCAGA ACATATTGTT | 2760 |
|    | CGAGCCATTA GAAAGAGACA AACGATTTTG TTATATGCCG TAACAGGTGC AGGTAAAGCA  | 2820 |
| 55 |  |      |

|    |            |            |            |             |             |             |      |
|----|------------|------------|------------|-------------|-------------|-------------|------|
|    | TCACCAGGTG | TAGATGTTGT | TGTAGAAATT | AGTAAACGTA  | TTAAAGACGC  | ATTTCTTAAT  | 2940 |
| 5  | GAAGATATAG | ACATACTACA | CCAGCAATCA | AGACAAACAT  | TTGAAGGGCA  | TTTTGTGTGA  | 3000 |
|    | TGCACAGTGC | ATCAACTTTA | CCGATTCAAA | CAGCACTTTG  | ATACTATTTT  | TATTGATGAA  | 3060 |
|    | GTGATGCCT  | TTCCCTTTAT | TAATGTAAAC | AAGCATTGAA  | GTCACTTTCT  |             | 3120 |
| 10 | AAAGTTGAAC | ATGCATCAAT | TTATATAGTA | GCAACACCAC  | CGAAACCTTC  | TCTGTCCAGAG | 3180 |
|    | ATTGCCACAG | AAAAATAAAT | TAAATTGCCA | GCTCGCTTTC  | ATAAAAAATC  | ACTTCCAGTT  | 3240 |
|    | CCTAAATATC | GTTATTTCAA | ACTTAATAAT | AAGAAGATTG  | AGAAAAATGT  | ATACCGAATT  | 3300 |
| 15 | TTACAAGATC | AAATTAATAA | TCAACGTTAT | ACACTGGTGT  | TTTTTAAACA  | TATAGAAACA  | 3360 |
|    | ATGATTAAAA | CATTTTGGGT | TTATAAGCAG | AAAAATCTACT | AATTAAACATA | CGTCCATAGC  | 3420 |
|    | GAGGATGTTT | TTCCGCTTGA | AAAAGTTGAA | CAATTAAGGA  | ATGGACATTT  | CGATGTCATT  | 3480 |
| 20 | TTTACTACGA | CAATATTAGA | ACGTGGATTT | ACAAATGCCAA | ATTTGGATGT  | TGTTGTTATC  | 3540 |
|    | GATGCACATC | AATATACTCA | AGAGGCTTTA | ATAGGTTTAC  | CTGGACGTG   |             | 3599 |

(2) INFORMATION FOR SEO ID NO: 266:

25 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1017 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 266:

|    |            |            |            |            |            |             |     |
|----|------------|------------|------------|------------|------------|-------------|-----|
| 35 | TTTCACAGAA | GGCTTgAAAA | AtGTTwCAG  | TGGTGCcAmC | CCAGTTGGTT | TACGACAAAG  | 60  |
|    | TATCGACAAA | CGAGTTAAAG | TGCTGTGTGA | AGCGTTACAT | GAATATTCTC | AAAAAGTTGA  | 120 |
|    | AAATAAAAAT | GAAATTGCGC | AAGTAGGTGC | GATTTCAGCA | GCAGATGAAG | AAATTGcAGC  | 180 |
| 40 | TTATATTCT  | GAAGCTATGG | AAAAAGTtAG | TAAcGATGGT | GTcATTACAA | TTGAAGAATC  | 240 |
|    | AAATGGACTA | AACACTGAAC | TAGAAGTGGT | TGAAGGTATG | CAATTtGATC | GTGGTTATCA  | 300 |
|    | ATCACCgTAT | ATGGTTACTG | ATTcAGATAA | AATGGtTGCT | GAATTAGAAG | GCCcATACAT  | 360 |
| 45 | TTTAGTAACA | GATAAGAAAA | TCTCGTCTTT | CCAAGATATC | TTACCTTTAT | TAGAACAAGT  | 420 |
|    | GGTTCAATCT | AATCGTCCAA | TCTTAAATGT | AGCTGATGAA | GGTGAAGGCG | ATGCATTAAc  | 480 |
|    | AAATATCTCG | CTAAACCGTA | TGCGTGGCAG | ATTTCACAGT | GTTCAGTCAc | AAGCACACTG  | 540 |
| 50 | TTTTGGTGAT | CGTAGAAAAG | CGATGCTtGA | ATTtTAGCT  | ATTTTAACTG | GTGGCGCAAGT | 600 |
|    | GATTACTGAT | GATTTAGGCT | TAGATTtTAA | AGATGcATCA | ATTGATATGT | TGTGACTCTGC | 660 |

CAGCATGAT GCACGTGTTA GCCAATTGAA ATCTCAAAT GAAGAACTG AATCTGACTT 780  
 TGATCGTAA AAATTACAAG AGCGCTTAGC TAAATTAGCA GGTGGTGTG CAGTTATCAA 840  
 5 AgTAGGTGCA GCAAGTGAA CAGAGCTTAA AGAACGTAAA TTACGTATTG AAGATGCATT 900  
 AAATTCACA GTGCAGCAG TTGAAGAAGG TATTGTTGCA GGTGGTGGTA CTGCATTAGT 960  
 10 AAATGTTTAC CAAAAAGTAA CTGAAATTGA AGCTGAAGGT GACATTGAAA CAGGTG 1017  
 (2) INFORMATION FOR SEQ ID NO: 267:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1409 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear  
 15  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 267:  
 20 TTAATCCAGC GTTAACTGTA TTTGCATTTA TTATGATTAT TTCGATTCTT TTAGCGTATG 60  
 TATTTAAATG GCTTGGATTA GTGGATGATG TGTATTAAAT GGTCAATTATC ATTTCAACTA 120  
 25 TTTCCTTAGG CGTAGTTGTT CCAACTTTAA AAGAAATGAA TATTATGAGA ACAACTATAG 180  
 GGCAATTAT CCTATTAGTA GCAGTACTTG CGGACTTAGT AACTATGATT TTATTAACGG 240  
 TCTATGGCGC AATCAATGGT CAAGGCGGCA GTACAATATG GTTAATAGGT ATATTAGTTG 300  
 30 TTTTCACAGC AATTTCATAT ATTTTAGGTG TTCAATTAA AAGAATGTCA TTTTACAAA 360  
 AATTAGTGA TGGTACGACG CAAATCGGTA TTGTCGGT ATTTGCATT TAATATTAT 420  
 TAGTAGCCCT AGCAGAGGGA GTTGGCGCAG AAAATATATT AGGTGCAITC TTAGCAGGTG 480  
 35 TCGTTGTTTC ATTATTAAAT CCAGATGAAG AAATGGTTGA AAAGTTAGAC TCATTGGTT 540  
 ATGGGTTCTT TATTCCTATT TTCTTTATAA TGGTGGTGT AGATTTAAAC ATACCTTCAT 600  
 40 TAATTAAAGA ACCGAAATTA CTAATTATCA TACCGATTTT AATCGTnGCA TTTATCATT 660  
 CAAAATTAAT TCCAGTCATG TTTATTCGAC GTTGGTTTGA TATGAAACCA ACGATTGCAT 720  
 CAGCAATTTT ATTAACATCA ACATTATCGC TCGTGATAGC TGCAGCCAAA ATTTTCAGAAA 780  
 45 GATTAAATGC TATTTCAGCT GAAACGTCAG GTATATTAAT TTTAAGCGCA GTCAATACAT 840  
 GTGTATTCTG TCGATTATT TTCAAAAAC TGTTTCCAGT TCCAGATGAG TTTAACCGTA 900  
 AAATTGAAGT TAGTTTAATT GGTAAAAATC AATTAAACGAT TCCTATAGCG CAAAATTTAA 960  
 50 CATCTCAGTT ATATGACGTG ACATTATATT ATCGCAAAGA CTTGAGTGAT CGTCGTCAT 1020  
 TGTCAGATGA TATCAGCATG ATAGAAATTG CTGATTATGA ACAAGATGTT TTAGAACGAC 1080  
 55

|    |  |      |
|----|--|------|
|    | AAGTGTCTAA ATTAGCCAAA GCACATCAAG TTGAGCGTGT CATTTCGAGA CTTGAAAGCA  | 1200 |
|    | CAACGGAGCA TACAGAGTTA GTTGATTGAG GTATTGAAAT TTTCAGTAGC TACTTAAGTA  | 1260 |
| 5  | ATAAAATCTT ATTAAGAGGT TTAATTGAAA CACCTAACAT GTTGAATTTA TTAAGTAATG  | 1320 |
|    | TTGAAACGTC ACTATATGAA ATTCAATGT TAAATTATA ATATGAAAAT ATTCAATTAC    | 1380 |
| 10 | GTAAATTTCCC ATTCCGAGGA GACATCATC                                   | 1409 |
|    | (2) INFORMATION FOR SEQ ID NO: 268:                                |      |
|    | (i) SEQUENCE CHARACTERISTICS:                                      |      |
|    | (A) LENGTH: 4702 base pairs  |      |
| 15 | (B) TYPE: nucleic acid   |      |
|    | (C) STRANDEDNESS: double   |      |
|    | (D) TOPOLOGY: linear   |      |
|    | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 268:                         |      |
| 20 | AAAGAGGGTT TTTTACTACT ATAATCCATC CTTTAATGGA ATTCCATTG TCCCTTTGGC   | 50   |
|    | CCGTTCATAA CCATAGGAAA ATATATGChG hAATCATTGA TAGTAAATAC ATCAATAATA  | 120  |
| 25 | CGAATAATGT CGGTGAAATC ACGTCCGTAT TACCATTATT AGCTAAAACA TTTTCCAAGT  | 180  |
|    | TTTCTTTTGA ACCGGATACT CATAAATGCT TTTAATGChT GGTTTTGCTC GTCGCCATCT  | 240  |
|    | TTAGAATTGG TAATACTTTT TTAAATTCCT TAGCGAACAA CTCATTATCT TTATCGTTTT  | 300  |
| 30 | TAGCCATTGG ACGATTCAAA TCGTTAGCTC TTACGCCCTAA TGCTCGACCT GTCGCATGTA | 360  |
|    | GCTTATTGAC ATTATTGTG ACTTGATTTA ATTGGCCACT AACAGAATCT GCAATTGATT   | 420  |
|    | TTGATTCTTG TGTATCTGAT AGCAATTGCG TACTCTTTTC AGAAAATCTA CTAATTCTT   | 480  |
| 35 | TATCTAAATT TGAAGACATC GTATTAAATT CATCATTTTT GCCTTTATCA ATTTTGGTT   | 540  |
|    | CTTGTGGCTC TTCAGCAAAA GTCTTTTTAA CGTTTTCTAA CTGATCAATC AGCTTGGAAA  | 600  |
| 40 | TATCTCTTK ATTTTGTGTT gTATTCTTTT TGTTATTAA AATGTCATCA ATCAGTTTGT    | 660  |
|    | CTGAGTTTTC TTCCATTGAA TCAATTGAT GTAACACAGC TACTTTATCG TCTTTGAAAC   | 720  |
|    | TTTCCATGTC ATTGATAACT TGGTCAACCA TCATATCAAT TAAACGTTTG TTGTGCAATG  | 780  |
| 45 | GTTTATCTTC tCTGCCTTTT GTATCTGTGT ACATTTTATA ATGCGCATCA AACCAGATA   | 840  |
|    | ATGCACTCAA TTGCTGGCTT AATGCATCTT TCGATAAGCG ACCATCaAGG TTATGATTCA  | 900  |
|    | ATGTACATC CACAACACTC GTTGCTTTCT TATCATTTGG TTCACTTGA CGATTGCTT     | 960  |
| 50 | GTCCAAATAA CAATTGTAAA TGCACTGTTT TATCTTTTAA GAAATCTTTC TCAGCATCCT  | 1020 |
|    | TTTTCAATTG AGCAAGCCCA TTGACTTCAA CTTTATATTC CTTGTAGAT GTATCGAGTT   | 1080 |



EP 0 786 519 A2

|    |   |      |
|----|---|------|
|    | TAAAAATGAGG ATCTGTTGCA ACAGTTAAAT GATTAATAATC ATTACTTTTA ATCGTTTCAG | 1200 |
|    | TACGTTGCAC TTTGACACCA TCATTAAATCA AACTACTTGT GTCTTGCACA ACTTGTTTAT  | 1260 |
| 5  | CGTAATCTGT TAAATTAATG TGTCTGCTA ACGGTTTTTT CAAATTATAT TCATTTTTAT    | 1320 |
|    | AAGTTTTTGC TTCCTTGACA ATGCTTCTGT ATTTATTAGC TTCATCCTCA TTTAAACCTG   | 1380 |
|    | CAGCTATAAA GTCTTGTTTA GACATGTTAT AGATAAATGT TGTATCTGTA TCAGGTTCTT   | 1440 |
| 10 | TGACAATATC ATCATGAAGT TGTCTCTCTA AGTTTTCAGC GAATTGAGCA TTGTTCAATT   | 1500 |
|    | TAATGCTATT TAGCGCATCT TGTAAGTCTT TGTATTGTC AAGCTCATCT TGCAGTGATT    | 1560 |
| 15 | CTGTTAATTG CTTACGATAT TCTTCAATCA TACCTTTTGA AAATGGTGAC TCTTGTGATT   | 1620 |
|    | GAATGATTC TCTTAATTTA TCTAAGTTTT CTTTAACAGT TTGTTTATAT TCTTCTTTAC    | 1680 |
|    | CTGTATCTTG CATACTTGAT TGTGTATCAA TTGGCTGTC CATCTGTTTT AATGCATTGA    | 1740 |
| 20 | TATAGTTATC AAGTCCACG CTATCTTTTT GCGATTATA ATCTTGTAAC ATTTTATCCA     | 1800 |
|    | TCGCTGTATT GTGCTGCTCA AATAATGAAT TTTGTTTTTC AATTAAAGTC GAAACATTAT   | 1860 |
|    | AATCTGTGTT CACTCTGAAC GTATCTGAAT TCGCACTCAA TAATGATTTA TTGTATGTT    | 1920 |
| 25 | GGAACCAATT TGTAATGTCT TTGTTTGAGC AAATTTGAAT TACAAGCGTA TCTGTAAATA   | 1980 |
|    | ATTCGGGAA GTCGTTAATT GGATTTAATA AGTAATTCGA GAATTTACTA TTCACCCAT     | 2040 |
|    | GTTACGCGT CATATAGCG CCAACATTTT TTGTGCAAT ATGTAAATTA TCAATGATCT      | 2100 |
| 30 | TTGTAAATA AATTTCGACT AAGTTTTTGT TAAAGTCGTT AAGTACATTA CTTACAACCT    | 2160 |
|    | TTTCTGTGTT TTAGCTACT TCTTCTTTTT GTCTACAGC TGTTTTATAC TGTAGCGATA     | 2220 |
|    | TTTTCGATGG TGTTTTAGCG TCTAATTGCA TTGCCAATTT TGAAAAGTTT TCTGGGATAA   | 2280 |
| 35 | CAATCATGAC TTGGTATCCA CCATTTTTC AACCAGACTC AGCAACGTTT CTGTACTG      | 2340 |
|    | TTTCAAAATT ATAGTTTTTC TCATTTGCTA ACCTTTTAAAT AAATGCTTGA CCCAGCTCAA  | 2400 |
| 40 | CTTTTTTACC GTTATATGTC GTTGGTTGAT CCTCGTTAAC AATTGCGATA TGTATTTTAT   | 2460 |
|    | TATTTTATT ACTTACACTT TGGGATCCTT TTCTGTATTG ATCTCCATAT TTTGTTTGAA    | 2520 |
|    | CAAAAAATAT CATACTAACT ATGGCAATTA TAATAATTAA AGTGACAATT AATGCATAAA   | 2580 |
| 45 | TCCAATTTTT CTTTTTCATG CTTATTTCCT TTCAGTGTG TTCTTAAAAA AATGATAAGC    | 2640 |
|    | AAAGCCACAT TAGAAAATGT GACTTTGCCA ATTTCAGAAT GCTTATTGCA AACCGAAATT   | 2700 |
|    | ATTAGAAAGT TGTGTGCTT GTTCTTGAAC GGCATCAGCA GTGCTATTGA ATTTGTTGTT    | 2760 |
| 50 | AATTTCTTCT AATAATTGTG CAAATTTTTC TACTTTAGGA CTAAGTGTG GGAATTGCTC    | 2820 |
|    | TTCGAAACGG CTGAAGCTT GACCTTCCCA GTTCGCTGCA ATTTCACTCT GTGCACGKGT    | 2880 |

55

|    |             |             |            |            |             |             |      |
|----|-------------|-------------|------------|------------|-------------|-------------|------|
|    | TCGTATTCC   | TCGTGACTCA  | TCTTAATCAT | TGCCATAACT | AAGAACTCC   | TGAATATTT   | 3000 |
|    | AAGTTTATCA  | AAACTTTTTA  | GGGACACTAT | TTTTTGAAAA | AGTGTCTCTT  | ACTCAAAATA  | 3060 |
| 5  | TATATAAATT  | ATTAGTATAT  | GTATATAGT  | TTTTAAGTAT | TTTTAGCTTT  | TTTAAAAATA  | 3120 |
|    | ATATATTGAA  | TATAACCATA  | TATTTTAAAT | TAACCATTCA | TTTTTGTAAT  | ATAAATGTGT  | 3180 |
|    | ATACTAAAA   | TAAATTAAT   | ACATAAAGGA | TTAATGGT   | ATTATGAAGA  | AAACAATTTT  | 3240 |
| 10 | ACTGACGATG  | ACAACTCTTA  | CTTTATTTAG | TATGTGCGCT | AACTCGGCTC  | AAGCATATAC  | 3300 |
|    | GAATGATAGC  | AAAACATTAG  | AAGAAGCAAA | GAAAGCACAC | CCAAACGCAC  | AGTTCAAAGT  | 3360 |
|    | GAATAAGAC   | ACCGGCGCGT  | ATACTTATAC | ATATGACAAA | AAACAACACG  | CAACAACAAA  | 3420 |
| 15 | TCATCAAAAC  | CAGTCACGTA  | CAACGCACAA | TCATCAACAC | CGAAATCAAC  | GTGATCTTAA  | 3480 |
|    | CAACAATCAG  | TACCATCTCT  | CATTAAAGTG | TCAGTATAGC | ACGCAATTAAG | ACGCAATTTGA | 3540 |
| 20 | TTCACACACA  | CGCGCTCAAA  | CGTACCAAG  | CAATCCTTTG | ACACCAGCAA  | TACCGAATGT  | 3600 |
|    | CGAAGACAAT  | GACGATGAAT  | TAAATAACGC | TTTTTCAAAA | GATAACAAAG  | GGCTATTAC   | 3660 |
|    | AGGCATCGAT  | TTAGACGAAT  | TGTATGACGA | ATTACAAATG | GCCGAATTTA  | ATGACAAAGC  | 3720 |
| 25 | AAAGACCGCT  | GACGGTAAAC  | CTTTAGCATT | AGGTAACGGT | AAAATCATTG  | ATCAGCCTCT  | 3780 |
|    | TATCACAAGT  | AAGAACAAC   | TATATACTGC | TGGCAATGT  | ACATGGTATG  | TCTTTGATAA  | 3840 |
|    | ACGTGCCAAA  | GATGACACAA  | CGATTAGTAC | ATTTTGGGGA | GATGCTAAAA  | ACTGGGCAGG  | 3900 |
| 30 | CCAAGCTTCA  | AGCAATGGCT  | TCAAAGTAGA | TAGACACCCA | ACACAGGGAT  | CAATTTTACA  | 3960 |
|    | AACAGTAAAT  | GGTCCATTGG  | GTCATGTAGC | CTACGTTGTA | AAAGTTAATA  | TTGATGGAAG  | 4020 |
|    | TATTCTAATT  | TCAGAAATGA  | ACTGGATTGG | TGAATATATC | GTTTCATCAA  | GAAACATCTC  | 4080 |
| 35 | TGCTTCAGAA  | GTTTCATCAT  | ATAATTACAT | CCATTAAAT  | AATCATGACA  | TCAATAAAAA  | 4140 |
|    | GGCACTCAGT  | CGCAGTTTAC  | AATTCGTAAC | ACTGCAAAAT | TGGTCGCTTT  | ATTTTGTATG  | 4200 |
| 40 | TTATTTCGATT | ATAAAATTAC  | AAGAANAATG | TCTCTACATT | CCCCATTAA   | CAAAATCGTT  | 4260 |
|    | TACGAAAGTA  | TAAATGTAGC  | TATAAATAT  | CAAGTCGTAA | CAACTATGGG  | CATCATCGTC  | 4320 |
|    | TTGAATAAGA  | ATACACGCTA  | TTTTTCTCTG | GCATATATAT | CCAGTACTAG  | CCAAATTAATA | 4380 |
| 45 | ATGATTATAA  | CAACAACAAA  | ATAAATAACA | GGATTTCATG | ATATAGCATC  | TGCGTGAAC   | 4440 |
|    | TCAGGTGTCG  | TTCTTAATTT  | AGTGAATAAT | AACATCACTA | CTGAAATAAT  | GAAAAAGTAG  | 4500 |
|    | ATACCTCTTA  | TCTTTGATGT  | CTGTAATATC | AATTCCTGCT | CTCATATGAC  | CTCTTTAGAT  | 4560 |
| 50 | TCACCCAATT  | CTTTTGAAT   | CAATAAATTT | ACTACCTTAG | GTTTCACCCA  | TAAACACTTA  | 4620 |
|    | ATTGCAAAGT  | ACATAAAAAAT | ATATGATCCA | GTATCCATAA | ACATTAAAAA  | GTTGCTTAAA  | 4680 |

## (2) INFORMATION FOR SEQ ID NO: 269:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2004 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 269:

|             |             |            |            |            |             |      |
|-------------|-------------|------------|------------|------------|-------------|------|
| ACAAAAAATT  | CACCCCTCATT | AATATTGAAA | CTAATGTTAT | CGACAGCAAC | ATGTTTGCCA  | 60   |
| TAACGCTTAG  | TTACATTTTT  | AAACTTAATC | ACTTTGCCAC | CTCTTTTTTT | CTCATAGCAT  | 120  |
| AAAACCGAGA  | TTATATGTAT  | GTATTCCTA  | TTTAACCAAG | TTTATTACAA | TTTTCAAATT  | 180  |
| TAAATGATTT  | ATCCTTGAAC  | TTTTTAATA  | AAATAATGAA | TAAWAGGWAA | TCWCCAGTTA  | 240  |
| AGAAATAGTG  | TTATTTTACC  | TTGAATTCAA | AAAAACACCC | AGTAAACCAA | GGAATGCTTA  | 300  |
| CTAGGTGTCT  | TCACTATACT  | TTGGCTTTAT | AATTTTGAAT | CGTTCTTAAA | AATGCTGGAC  | 360  |
| AATAATGTTT  | TAATTTGTAA  | CTACCTACGC | CATCATATTT | AATCATATCT | TGTTTCGAAG  | 420  |
| CAGGCTTACG  | TTTAGCAAAAT | TCCTCCAACG | TGTAATCAGA | AAATATACTT | ACAGGTGCTA  | 480  |
| TGTTAATTTT  | GTCACTTAAC  | TTTTTACGAA | CTCTACCAA  | CTGACTGAAT | AATACTCGGT  | 540  |
| CAACCCCTTC  | AACCGTATTT  | ATAAATACCT | TTTCAGTCGC | TTTTTGCTTA | AATGGTGTG   | 600  |
| TGAATACTTC  | TACTTCATTA  | CTGAGTAATT | TTTTAATTGA | AGTATCACAC | ATTAATATTT  | 660  |
| CGTCATTTTC  | ATTTAAGAAC  | CCTTGAATC  | TTAATTCATC | TATTAAGTGA | CITTAATCTG  | 720  |
| ATGTTGTGTA  | ACCTTTCAIT  | AAACCATGGG | TTGAAATTTG | GTCAACCTT  | TTATACTTAA  | 780  |
| TATAATCTGk  | TGACTCTCCT  | CTTAACACTT | GAATGATAAC | ACTATAACTC | TCITGTGTGT  | 840  |
| TCATACGAGC  | GATGCAACTA  | ATAATCATCT | TAGCTTCTTG | TGTCATATTA | TATGATTTAT  | 900  |
| CTTGTGAAC   | ACAATTACTA  | CATGTGTCAC | ATTCTTCTAA | TTTTTCATTC | GGTTCAAAAT  | 960  |
| AATGGACAAAT | TGTTGCTTCT  | AGACATTTTT | TTGTTTTTGT | ATATTGAATC | ATTTTAGTTA  | 1020 |
| ACTTTTCGCC  | CATTTTATCT  | TTATAGTCAT | CATCAGCTTG | AGAGACTGTT | ATAAAATACT  | 1080 |
| CGTGTAAATT  | GATACGCGGT  | TCGTAAATA  | ACAAAATACA | TTCACTTTTT | AACCCGTCAC  | 1140 |
| GACCTGCAGC  | ACCCGCTTCT  | TGATAATAAG | ATTCTAAATC | TCCAGGCATA | TTATAATGAA  | 1200 |
| TAACAAGAGC  | TACATTGGAT  | TTATCAATAC | CCATACCAAA | AGCATTTGTA | GCAACGACTA  | 1260 |
| CTTTAACACG  | ATCAAAATAG  | AAATCATTCT | GCCTTCTTTC | TCTTTCTTTA | TTGCTCAAAAC | 1320 |
| CTGCATGATA  | TATAACACTT  | TCAATTTTCT | GACTTTCTAA | GGCTTCTTGA | AGCTCTTCAA  | 1380 |

EP 0 786 519 A2

|    |             |            |            |            |            |             |      |
|----|-------------|------------|------------|------------|------------|-------------|------|
|    | CCAAATATAAA | TTTTTGACGT | TGATAAGTAG | GATTTACTTT | AAAAATTAAG | TTTCTACGCT  | 1500 |
|    | TAGTACTCGT  | TTTAATTGGA | TCAGTTTGAG | CGATATTTAA | CTTTTCTCTA | ATATCTTGCT  | 1560 |
| 5  | GTACTTCAAC  | CGTGGCAGTT | GCTGTCAACG | CTATTATTGT | AAATCTTGA  | GGTAAACGTAA | 1620 |
|    | ATACTTTTGA  | AATAACATT  | TGGTAACTCG | GCCTGAAATC | ATGACCCCAT | TTAGAAATAC  | 1680 |
|    | AATGCGCTTC  | ATCAAAACGG | ACTAAGTGAA | TCTTTATACG | CTGAAGCAT  | TTTAAAAAAT  | 1740 |
| 10 | ATCGGTTTTC  | AAATCGTTCT | GGTGCAACAT | ACAAAAATTG | AATTTCTCCA | TTTGATAATG  | 1800 |
|    | CTTTTTCAA   | ACGTTGTTGC | TCTTTTGTAG | TCAAACTACT | ATTTAAAAAA | GCAGCTTGAA  | 1860 |
|    | TTCCCATCGC  | TTTAAATGA  | TCCACTTGAT | CTTTCATTAA | TGATATTAGT | GGACTTATTA  | 1920 |
| 15 | CAATTGTTGT  | ACCACTTAAC | ATAAAACCTG | GTACTTGTA  | GCTATAGAC  | YTACCTCCAC  | 1980 |
|    | CAGTGGKGA   | GCACCAAGC  | ACAT       |            |            |             | 2004 |

(2) INFORMATION FOR SEQ ID NO: 270:

(i) SEQUENCE CHARACTERISTICS:

- SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 2244 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 270:

|    |             |            |             |            |            |            |     |
|----|-------------|------------|-------------|------------|------------|------------|-----|
| 30 | AAGATTGCT   | TGCGTTGAGG | GTTTATATAT  | CTGACTCAAT | TGCCACATTT | TTATCAAGAG | 60  |
|    | TAGTGTGATA  | TACTCATCAT | AATTATAGCT  | AATATTATAT | TTTTTTAAAA | GATAGTGTAT | 120 |
| 35 | GATTTTCIGG  | TGTTTGTGT  | ATACGTCATT  | AAATTTCAAG | TAGTCATTCT | CCAAGTTATA | 180 |
|    | CGTATAACAA  | ATATTTCCGG | ATAAAGTTAG  | ATAAAAAAT  | TTAGAAAAAT | CATTCTATTG | 240 |
|    | CGTATCGCT   | AAATTAAGTG | TAAATATAGA  | GACATAAGTA | ATTAATTTAA | TGTGATATGA | 300 |
| 40 | TGTATTATTA  | CTTTGCTAAA | TGTAGATAGT  | AACAAAAATT | GTAATCGGGA | GGTAACAATG | 360 |
|    | GATTACGCAC  | ATTTAAATTT | AGAACATTTT  | TTTGACAGAA | ACGACGATT  | AGATGTTATA | 420 |
|    | AGAGATCGCG  | CTGATTTTCG | GATGATAAAT  | AACCTCACTA | ATGAAATGAT | GATACGTGAT | 480 |
| 45 | GGTCAAAATG  | AAGGCAGAGT | TGATTTAAAT  | CAGTACTATT | ATAAAAATAG | ATCAAAATGC | 540 |
|    | GCAAGTTTTA  | TTATGATGGA | TTATAAAAAA  | GAAACTAAGT | AAACGAACAA | AAGAATTTTT | 600 |
|    | TGTTTTTTAA  | TACGTGAATA | ATAAAGATTAT | TGATATAAAG | GTTTTTCAAG | GTTATACAAA | 660 |
| 50 | AAAGATAAAAC | ATTTATGATT | CGTAGATCAA  | CGTAAAGTAA | TGTTGTAAAA | TGGTTTAAAA | 720 |
|    | CGTTTCATTT  | ACATTACTGT | TTATTATGAT  | TGATGTAAAC | ATGCATAGAT | AAATTTGGTA | 780 |



ATAAGCAATT TAATTTTGAG TCTACAATGG AGGAATTATC ATCITTATCA GAGACTTGCC 60  
 AACTTGAAGT GTTGGGTCAA ATTACTCAAA ACAGAGATCG TGTAGATCGC AAATATTATG 120  
 5 TTGGTAAAGG TAAAAATTGAA GAAATTCAAG CATTATTTGA GTTCAAAGAT ATTGATGTAG 180  
 TCATACAAA TGATGAATTA ACGACTGCAC AATCCAAATC ACTAAATGAA GCTTTAGGTG 240  
 TAAAAATTAT TGATAGAACT CAGTTGATTC TTGAAATATT TGCATTAGA GCAAGAAAGTA 300  
 10 AAGAAGGTAA ATTGCAAGTA GAGCTAGCAC AACTTGATTA TTTATTACCT AGATTGCAAG 360  
 GCCATGGTAA AAGCCTTTCT CGTTTAGGGT GCGGTATTGG AACTAGAGGC CCTGGTGAAA 420  
 CGAAGTTAGA GATGGATCGC AGACATATTC GAACTCGTAT GAATGAAATT AAACATCAAT 480  
 15 TCGCGACGGT AGAAGAACAT CGCGAAAGAT ATCGAAATAA AAGAAATCAA AATCAGGTGT 540  
 TTCAAGTAGC TTTAGTTGGT TATACAAATG CTGGTAAATC ATCATGGTTT AATGTTTTAG 600  
 CAAATGAAGA GACGTATGAA AAGATCAAT TATTGCAAC GTTAGATCCT AAAACACGAC 660  
 20 AAATTCAAAT AAATGATGGA TTTAATTTAA TTATTCAGA TACTGTTGGT TTTATACAGA 720  
 AACTACCTAC GACGTTAATT GCAGCTTTTA AATCAACTTT AGAAGAGGCT AAAGGTGCAG 780  
 25 ATTTATTAGT ACATGTCGTA GATAGTAGCC ATCCTGAATA CCGTACGCAG TATGACACAG 840  
 TTAATGATTT AATCAACAA TTAGATATGA GTCATATTTT TCAATAGTT ATTTTAAATA 900  
 AAAAGGACTT ATGTGATCAT GCATCAAATC GTCCAGCAAG TGATTGCGCT AATGTTTTTG 960  
 30 TTTCTCTAA AATGATGGT GATAAATTAC TTGTTAAGAC GTTATTTATT GATGAAATCA 1020  
 AAAGGCAATT AACTTATTAT GATGAGACAA TTGCGACGAA TAATGCAGAT CGATTATATT 1080  
 TTCTAAAACA ACATACATTA GTGACTGAAC TTAATATGTA TGAAATTGAA AATGTTTATC 1140  
 35 GTATAAAGG ATTTAAAAAA TAATAAAGG ACGAAATTCA AATGAAAGAT ATAAGTAAGA 1200  
 TAGTAGCTGA CGTCGAATCA ACGTTAGCAC CATATTTTAA AGAAATTGAA GAAACAGCAT 1260  
 40 ATATTAATCA AGAAAAAGTA TTAATGCAT TTCATCATGT CAAAGCAACC GAAAGTGATC 1320  
 TACAAGGATC AACAGATAC GGGTATGATG ACTTTGGACG TGATCATTTA G 1371

## (2) INFORMATION FOR SEQ ID NO: 272:

45 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 6035 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear  
 50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 272:

55

55

EP 0 786 519 A2

|    |  |      |
|----|--|------|
|    | ACCATGACAA TAGTGTATC AAACGATTIA TTCATTCTT CCAAACGTTG TAATAGGTCA    | 1920 |
|    | TTTGCACTTT TCGAGTCGAG TCGCGCTGTT GGCTCATCTG CAAATATGAT TTGTGGTTTG  | 1980 |
| 5  | TGAACAAATG CTCTCGCTGC TGCAGTTCTT TGTGTGTGAC CACCAGATAA TTCGCTAGGG  | 2040 |
|    | TATTTATTTT CTAGGTCATA AATACCTAAT GCTGTGTGTA TCGCTTTATA ATTTCTTCTC  | 2100 |
|    | ATTGTTGCCT TCGACATTTT TTGAACAGAT AAAGGTAACA TAATGTTTTT TTTAACGGTT  | 2160 |
| 10 | AATGTCGGCA GAATACTGTA ATCTTGGAAG ATGAAACCTA ATGATTCTTT GCGGAATTTG  | 2220 |
|    | GCAAGTGGCT TTTGATTAAAG TTTATTAAAG TCTTGCTCGT TAGCAATCcg cTACCGCTAG | 2280 |
| 15 | AAATTGGTCT AATTGAACTT AGTACATTTA ATAAGGTTGT CTTACCTGAT CCAGAAGGCC  | 2340 |
|    | CCATAATCGC AAGCAATTGC CCTTTTTGTA TGTCAAAGTT AATATCTTTA AGTGCTTGAA  | 2400 |
|    | ATGTGTGCTT TTTACCGTAT GTTTTTGAAA CATGTgCACT GATAATATCG TCATAGTCTC  | 2460 |
| 20 | ACTCCTTtTG TAITTAATTT CATTTTAAAT AATGTTTGGG GTAGTAGCCT TTATCTAAAC  | 2520 |
|    | TTACAATTCa ATGAATGAAC CTTACAGAGT TGAARcTAT CGCTACTTAG TAGATTTTTG   | 2580 |
|    | AGTGAGGATA CAGATTCATC GTACATATTA GACAAAAGCA ATGGTGCTTT CTAAGTGATG  | 2640 |
| 25 | ATGTTTGTGT AAATTGAGAA AAGGGAATTT AATTATTGTA TAATAAATTT TTTGTAAAAA  | 2700 |
|    | TTAAAAGAGG GTTTTATTGT AAAGGAATTG ATTGTTATGG AAAAAGGAAA TCAAGGTATT  | 2760 |
|    | AAATGGTCTA GTTTAATAAT GGGTGATTA TTAITTAATGT TGGCAGTCGT TATTTTACA   | 2820 |
| 30 | TTTCCAATTG AAAATTTTTA TGCTATTACC TGGTTGATTG GACTGTTTGT ATTAATTTAAC | 2880 |
|    | GGTGTGATTC AAATCGTTTA CCGTAGAAAA GCAAAAAGCTT TAGTAGGTGG TAACCAAAAT | 2940 |
|    | TGGATTCTGT TTAATGGGAT TGTAGATATT CTATTGGTTC TATTAGTTAT TTTTAATGTT  | 3000 |
| 35 | GGCGCAAGTT CAGCAATCTT TATTTATATG TTIGCTTTTT GGTTTATTTT TAGTTCATATC | 3060 |
|    | TCTGGATTAT TTACGTTTTT GGGTAGTGGT AGCTTAAAA TAATTTCACT GATTTTAAAT   | 3120 |
| 40 | TTATTAGGTA TTGTTTTCGG TGTCAITTTA TTATTTAATC CATTAATGGG TATCGTCTTT  | 3180 |
|    | ATTTCGACGA TGATTGCTAT TGCATTGTGA TTCGTAGGTG TCATTTATGT TGTAGATGCA  | 3240 |
|    | CTTGCTTAAG TAAAATGAAG CCGTTCAAAA GAAGGGTGTG ACATGAAGTT TGTGTATAT   | 3300 |
| 45 | CCTTTTGTGT GTGTTTATGA AGCATAAAAA AGGGGGGCTA CCTACAATAA GTAAGATACG  | 3360 |
|    | CCCATATTTT TATATTTTAC TATTATTGTT TTTCAATACG ATTAATAGTT ACATTTAGTC  | 3420 |
|    | CAAAATATTT TTCTAAAAAA TGTTTATAGT TATCTTTAGT GACATCAAAT TTTTCTGAGC  | 3480 |
| 50 | TACCATTCCT TGTTAAAGTT AAATGATTTT CAGACATTGT AGCACGGCCA AATGATTGTG  | 3540 |
|    | GCATTGTAAT TAATAAATGC TGTACAAATA TTGAATCTGG ATGCGTTTGA TTATATTCGA  | 3600 |

55



EP 0 786 519 A2

|    |             |             |             |             |            |             |      |
|----|-------------|-------------|-------------|-------------|------------|-------------|------|
|    | AATGATCATT  | TCGAATTTT   | TGAACATAGA  | AAATATCCTT  | GTCTTCGTTG | TTAAAAATAG  | 3720 |
|    | CGCGGAATGT  | ACCACTGATA  | TCAGTAATTG  | GTTGTGTATG  | CTCAGATGAA | GTAATAGGAA  | 3780 |
| 5  | TGGCATGTAG  | AGGTAAGTCT  | CCAAAGCCAA  | CATCAGTTAC  | ATAGAATACA | TCATTTATAG  | 3840 |
|    | AAACAACAAG  | TGAAGCATGT  | GAACCGTTCA  | GACTACGACC  | GCCACCGGGW | GTGTGAATAG  | 3900 |
|    | TAGCTGCACAT | TAATTCAGGA  | TTAAATCCTT  | TTTGTGTAA   | ATAGGCTTTG | AAAAATGTAT  | 3960 |
| 10 | TTAATTCATA  | ACAAAAACCA  | CCACGTTTAT  | CATGAACAAT  | TTTATTAAAA | AGTGCATCGA  | 4020 |
|    | TATTTATAGA  | TATGGGCTTA  | CTATTTTGAA  | CATCAATATT  | TTCAAAAGGT | ACAGTTAAACA | 4080 |
|    | TAAAACGTGT  | TGCATAATAA  | TTTAATGCTT  | CAATACTCGG  | TCGATTATAA | CGAGATGAAT  | 4140 |
| 15 | CAATTTGTAA  | ATAATTTCTCT | AACCTCGCAA  | TATTCATAAG  | CATAGCGCCT | CCTGTATTAA  | 4200 |
|    | AGATTATAAT  | TAAATTTTAA  | ACAGAAATAC  | TGAAATTTTA  | AATTCGAAAG | CATTGAATTT  | 4260 |
| 20 | TGGATAAATA  | CATTTTAAAT  | AGAAAAATAC  | GCTCTCAAAA  | TGAAGTCATC | TCTAAAAGAA  | 4320 |
|    | ACGATTTAAA  | GATGACTACT  | GAGAGCGTAG  | CATAATGGAA  | GAAGTGTGCA | GGGTGTCTAA  | 4380 |
|    | AAATGCAACA  | ATACAAAGGT  | AGTTGCAAGA  | CAAGTTCCTT  | TATCTAGACC | ATTTGTGTTT  | 4440 |
| 25 | TATGCGACCA  | AACCTCCAAA  | TTAAACTTGA  | AATAAGCCAA  | GTAATTAATA | ATAATGCAAC  | 4500 |
|    | TAAATATAG   | CCTAAATAAT  | CAAAATTCGAT | CGAACCAATG  | AATGCCCAAA | ACGCAACCATG | 4560 |
|    | TAAATCTAAC  | TTATCAGCAA  | GAATTTGTAG  | CAATTCAAATC | ATCCCAATCA | CTAATGCTGC  | 4620 |
| 30 | CATGACTGAT  | ATCGCAGTAA  | TCGTTATATT  | GTAATAGATT  | TTGCGAATAG | GATTGAAGAA  | 4680 |
|    | TGCCCAATTA  | TAGGCATACT  | TCATTACAAC  | ACCATCTAAT  | GTATCCAATA | AACTCATACC  | 4740 |
|    | TGATGCGAAT  | AAAATTGGTA  | AAGATAAGAT  | TCCGATAAAT  | GAAATGGCTT | GTTGTGATGC  | 4800 |
| 35 | GCCTGAAGAA  | AGAGCGAGTA  | ACGCAATTTT  | ACTAGCTGTA  | TCAAAACCAA | GTCCTCAATA  | 4860 |
|    | AAAGCCAAGT  | GGCAATACGT  | GCCAACTACG  | CGTGATTAAAT | TTGAAATAAG | GTCTCACAAA  | 4920 |
| 40 | TCGAGAAGAC  | AATCCTCTAG  | ATTCAAGTAA  | TGCATCGACT  | TCAGCTTCTT | CAATGTGTTT  | 4980 |
|    | ACGACGTAAT  | TTAGCGAACA  | AGTTAATTAA  | AGAGATTAAA  | ATAATTAGAT | TCAACACACC  | 5040 |
|    | GATAAGCACT  | AAAAAGAAAC  | CTGAAACTAG  | TGTACCAATC  | GTCCACCAA  | TATCTTGGAA  | 5100 |
| 45 | ATGCGGTAAT  | TCATCTTTAG  | CCCATTTTAC  | AGATACCCCT  | AAAAAACACG | CCATTAATAA  | 5160 |
|    | TACGACAGAT  | GAATGTCCAA  | TTGAAAAATA  | GAAAACCCACA | CCAGATGGAT | CTTTGCGTTG  | 5220 |
|    | CTGTAATAAT  | TTGCGAACCG  | TATTATCTAT  | TGCAGCAATG  | TGATCTGCAT | CAATGCGATG  | 5280 |
| 50 | ACGCAAACTT  | AATGTATATG  | CAAGAATCCC  | CATACCAAAAT | AAGATATGAT | GGTCTTTTCC  | 5340 |
|    | AGCAATCCAT  | AAAAAACTAA  | ACCCAATAAC  | GTGTAACAAA  | ATGACAATAG | CTATGTATGG  | 5400 |
| 55 |             |             |             |             |            |             |      |

ATATTTAATC ATACTGTATG TTCAATGGGC ACTCTAGTAA TAAGTGTCA TATAACAAAA 5520  
 ATGTATATGCC AAATTATTTG TTATATAAAA ATATACATGT AACACAAAA GATTTTITGC 5580  
 5 GATATATATA ATTTGATAAA TTAACCAACA ACAATGTAAG ATGTCACTTT GCTTAACCTG 5640  
 GCATCCTTTT TATGATTTTC AAATTCAAAA AAATGAGCAA AATGAATCTC TTAcCAGTT 5700  
 TTTAATATTT CaATACCATG CATGSAACCT AAGCACCCAT GTGTGATGCT GGAATGGATA 5760  
 10 TTGAGACTAG CAACCTGATT GTAATGATTA GATAGTTCTT GAATTAAAT TTGAGGTCCG 5820  
 TATATGTCAA AGCGCCAGG GACAGACCAA ATAAATCTG TTGTAACCAG TGAACGTAAT 5880  
 AATTCAATAT CTAATGCTGC TGTAACAACCT ATAAAACTA TCATTTGTTG ACGTTTAGGC 5940  
 15 GCATGATTGC ATGACACATC TCCTGTTAAC TTAAAAGGTA ATGATGACTG AACTTCCGTT 6000  
 TTAATAATGTA GTTGGTGCTG AATAAAGCT TGTTTC 6035

(2) INFORMATION FOR SEQ ID NO: 273:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1039 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 273:

30 TTTTGAACAG CCATATTTAT TCACCTCAA CATCATTATA ATGGTATTAG TCGCATTACC 60  
 TTCCTTGTT TTAGCTATAT ATGATTATAT GAGTTTAGA ATTTCTTCTG CTATTTTACA 120  
 ATTTCTAGGG GCTATCTCTT GGTCTTTTT ATCATTGATA TTATCGCTCA CACAATTTAC 180  
 35 ACCTTTTACA TTAGCGTCAT TTATAACTTC AATTATTTTG TTCACAAGCA CAATTATCAC 240  
 ATTAGCCATT GGTGTAAGT CTGTTGAAAA GAATGATTCC CCTTAAATTC CAAATGAAAA 300  
 40 AAAAGGTTCTG AAGGCCGCTA TAAAAACAG TTTTTCAGAA CCTCTATACT TCTATTCAAT 360  
 GATATATGGT TTGCAATTTT CTACCTTTAA ATCCACAGCT TCTGCCCTTG AAACCTTGTT 420  
 AAAATAAACC ATCAACAAC GAATGACAAC TTGATGTGCA ACAATGACAA TATCATCTTT 480  
 45 TTGTGTATCT TCATTGACAA CATGATTCAT AAAATGTTCT ACGCGTTGAT ATACATCTTC 540  
 ATAACCTTCT CCTTCAGGCG CTTTTTGTGA AAAACTATGA CGAAAGCTTT TAAAGTTTGG 600  
 ATCATTGAAA TATTTTTCAT ATTTGCGATT CGCACTGATT TCATCTTTAT ATTCAACCTC 660  
 50 AAATACGCCA AGTGAACGTT CTCTTAATAG AGGGGTAGTC GTTGATGCAA TGTCAATATGG 720  
 AAAAAATATGT TCAAACGTTT GCTGTGTTCT TAATAAGTCT GAAACATATA CATGTTTAAT 780



|    |  |      |
|----|--|------|
|    | TACCAAAATT ACTTGAAAGT GCAGATAATT TGATTAGTTC AACGAGTGCA CAATCTTATA  | 1260 |
|    | TCATTACAAA TGAAATGATA GAAAGAATTG CAGAAAATAG AAAGCAAGAT TCACTAGTAT  | 1320 |
| 5  | TGATTGATAT TGCAGTTCTT CGAGATATTG AACCTGGTAT TAGTGCCATC ACAAACATCT  | 1380 |
|    | TTAATTATGA TGTGTGATGAC TTAAGAAGTT TAGTTGATGC AAACCTACGT GAGCGACAAT | 1440 |
|    | TAGCGGCTGC AACAAATTCG GAACAAATTC CTACAGAAAT ACATGCACAC AATGAG      | 1496 |
| 10 | (2) INFORMATION FOR SEQ ID NO: 275:                                |      |
|    | (i) SEQUENCE CHARACTERISTICS:                                      |      |
|    | (A) LENGTH: 4826 base pairs  |      |
| 15 | (B) TYPE: nucleic acid   |      |
|    | (C) STRANDEDNESS: double   |      |
|    | (D) TOPOLOGY: linear   |      |
| 20 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 275:                         |      |
|    | CTTGATTTTT TCCCTTAGT ATTTTCCaTt TgantTGTGC AGCTTCTAAA TCCTGCTTTG   | 60   |
|    | GTTCTCTAGT GAACCTCATA ATTAAGCAG CTACAACGAA TGATACAAGT GCAGCAAGGA   | 120  |
| 25 | AGACACCGAG TAACATGTGC AAGAATTCAC CTCTAGGTGC ATTTAAACAG TAAACTATAA  | 180  |
|    | ATGAACCTGG TGACGCGGGA CTTTTAAATC CAAATCCTGT TGCTTGATAA GTTGCAACAC  | 240  |
|    | CAGTCATTCC ACCTAAAATA ACAGCGATAA ATAATAAAGG ACGCAITTAAT ACATATGGGA | 300  |
| 30 | AATAAATTTT ATGAATACCA CCTAAGAAGT GGATAATTTCC AGCACCATAT GACGTTGCTT | 360  |
|    | TTGCAGTGCC TTTTCAAAA ATCATATAAG CAAGTAAGAT ACCTAAACCT GGTCCAGGCT   | 420  |
|    | TAGATTCAT TGTGTATAAA ATTGATTGAC CAGCTTTTGC AGCTTGATCT GCACCAAGCG   | 480  |
| 35 | GTGTGAATAC ACCATGGTTA ATCGCATTGT TTAAAAATAC AATTTTGTGA GGCTCTACTA  | 540  |
|    | AAATTTCTTAC AAGTGAAGT AGGTGTGCAT GTACTAATGC TTCAACTGCC ATGATATAAA  | 600  |
| 40 | TATGCATAAT AAATTTTATA AGTGGTGCTA AAATTTTAAA TCCTGCAATC GTCATGATAA  | 660  |
|    | ATCCTAAAAA ACCAGCAGAA AAGTTATTAA ATAACATTTT AAAACCTTGC GGCCTTCTAG  | 720  |
|    | GTTGAATCAA TTGGTCGGTC TTCTTCATTA ACCAACCAAC AAGTGGACCC ATAATCATTG  | 780  |
| 45 | CACCAAGTAA CATTTGTTGA TCAGGTAATG CAACGATGAC CCCCATAGTT GCTGTTGCTG  | 840  |
|    | CGATGATACC ACCACGTAAA TCATAAATTA AACGACCACC ACTAAATGCG ATCAATAATG  | 900  |
|    | GGATTAAATA AGTAATCATT GGTCTGTGCTA AAGTAGCTAA ATCTTTGTGA GGTAACCATC | 960  |
| 50 | CATTATCTAT AAAAATGGCC GCGATAAAAC CCCAAGCGAT GAAAGCGCCA ATGTTTGGCA  | 1020 |
|    | TGATCATACT ACTTAAGAAT GATCCAAATG CTTGAACACG ACGACCAATT CCTTTTCTCT  | 1080 |
| 55 |  |      |

EP 0 786 519 A2

|    |            |             |             |            |             |             |      |
|----|------------|-------------|-------------|------------|-------------|-------------|------|
|    | GAGAGGTTAC | TTGTTACTCA  | ATATAAACAA  | AAATCAACTT | TGTCAAAATA  | AATGTGACAA  | 1200 |
|    | AATTAAATAA | AGTGTGATCA  | ATGTGACAGT  | ATAGATATTT | TGAAAAGTA   | AAACAAAAAA  | 1260 |
| 5  | ATTGTTTTAG | GATTTTAAAA  | ATTTTATTGT  | GAATAATTT  | GCAAAACAAA  | ACAAACCCGT  | 1320 |
|    | GTACAATAAT | GATTAATGGA  | AAGGGGGAAA  | GTTCCGCAGT | ACAGTTAAAG  | CGCCTGTGCA  | 1380 |
|    | AATAAATATT | TGTAATTTGAA | GATTAAGAGT  | TAATATATGA | GTGGCCTTTA  | TAGAGTGCAG  | 1440 |
| 10 | TATATGTATT | TGTAGACGAG  | GAGGATAGTG  | ATCGAATAGA | TCGGCGGAGT  | CTATCCCGCA  | 1500 |
|    | TGTGGCTCAT | TGTTTAGCTT  | ATTGAATAAA  | ACATTAGGGT | GACTTAATGG  | ACAAAGTTAA  | 1560 |
|    | TAGATTCGCC | AGAAAATTGAA | TAAAAAATA   | ATTAAATATG | AAAGTACAGT  | GTAGGCAATT  | 1620 |
| 15 | TGTATAGTTG | TAAAAATAAC  | TATGCTTAAT  | TGTTTATGGA | TGAATGCGAT  | GATAGCATGT  | 1680 |
|    | TCCTATTTAT | ATTATGAAAG  | CAGATTGTCA  | ATCTAAATTA | TCGGCAATAA  | ATCATAAATT  | 1740 |
| 20 | ACGCGTACTA | TTCCAATATG  | GAGGAAAAAT  | TCGTTATGTG | TGGAATTGTT  | GGTTATATTG  | 1800 |
|    | GCTATGaTAA | TGCCAAGAA   | TTATTATTAA  | AAGGTTTAGA | AAAAATTAGAA | TACAGAGGTT  | 1860 |
|    | ATGACTCTGC | AGGATATGCA  | GTAGTAAATG  | ATGATAATAC | AACCTGATTT  | AAAGAAAAAG  | 1920 |
| 25 | GTCGTATTGC | AGAAATTAGT  | AAAGTTGCTG  | ATAGTAGCGA | TTTGTATGGA  | CCTGTGGGAA  | 1980 |
|    | TCGGTCACAC | ACGTTGGGCA  | ACACACGGTG  | TACCGAATCA | TGAAAACCTCT | CATCCACATC  | 2040 |
|    | AATCATCAAA | TGGCCGTTTT  | ACTCTAGTTC  | ACACCGTGT  | TATTTAAAAA  | TATGAAGAGT  | 2100 |
| 30 | TAAAGGGTGA | ATACTTACAA  | TGCTATCATC  | TCACTTTTCA | AACAGATACA  | GAAGTATATG  | 2160 |
|    | TTCAATTAGT | TGAATTACTTT | TCAAATCAAG  | GACCTTCAAG | TGAAGAAGCA  | TTTACAAAAG  | 2220 |
|    | TTGTGTCAAT | ATTACATGGT  | TCATATGCAT  | TAGGTTTATT | AGATGCTGAA  | GACAAAGACA  | 2280 |
| 35 | CAATCTATGT | TGCTAAAAAT  | AAATCACCAT  | TATTTATTAG | TGTTGGTGAA  | GGTTTCAATG  | 2340 |
|    | TTATGCGATC | AGACGCACTT  | GCAATGTTAC  | AAGTGACAAG | CGAATATAAA  | GAAATCCATG  | 2400 |
| 40 | ACCATGAAAT | CGTTATTGTT  | AAAAAAGATG  | AAGTTATTAT | TAAAGATGCA  | GATGGAAGCG  | 2460 |
|    | TTGTAGAAGC | TGATTATAT   | ATTGCTGAAA  | TTGATGCATC | AGATGCTGAA  | AAAGGTGTTT  | 2520 |
|    | ATGCACTACT | CATGTTAAAA  | GAAATTCATG  | AACAACGACG | AGTAATGCGT  | CGTATTATTCT | 2580 |
| 45 | AAGAATATCA | AGATGCAGAA  | GGTAACCTGA  | AAATTGATCA | AGACATCATC  | AATGATGTTA  | 2640 |
|    | AAGAAGCAGA | CCGCATTATC  | GTTATTGTCAG | CAGGTACAAG | CTACCATGCA  | GGTTTAGTAG  | 2700 |
|    | GTAAAGAAAT | TTTAGAAAAA  | TGGGCTGGCG  | TACCAACTGA | AGTACACGTT  | GCATCAGAGT  | 2760 |
| 50 | TTGTCTACAA | CATGCCATTA  | TTATCTGAAA  | AACCAATTGT | CGTTTATATT  | TCTCAATCTAG | 2820 |
|    | TGAAACTCGT | AGATAGCCGCG | CCGCTATTAG  | TTGAAACTAA | TAAATTAGGT  | CTAAATTCAT  | 2880 |

EP 0 786 519 A2

|    |   |      |
|----|---|------|
|    | TACACGCGG TCCTGAAATC GCAGTTGCAT CTACAAAAGC ATATACTGCA CAAATTGCAG    | 3000 |
|    | TATTATCAAT CTTGTCTCAA ATCGTTGCAA AAGAGCAGG TCGTGAAGCA GATATTGATT    | 3060 |
| 5  | TATTGAGAGA ATTAGCAAAA GTAACAACAG CAATAGAAGC AATTGTTGAC GATGCACCAA   | 3120 |
|    | TTATGGAAAC AATTGCTACA GATTTCCTAG AAACAACAGC CAATGCATTC TTTATCGGAC   | 3180 |
|    | GTACTATTGA CTATAACGTA AGTTTGAAG GTGCGTTAAA ACTTAAAGAA ATTTCTTACA    | 3240 |
| 10 | TTCaAGCAGA AGGTTTTGCT GGTGGAGAAC TTAACATGG TACAATTGCC TTAATCGAAG    | 3300 |
|    | AAGGTACACC AGTTGTAGGT TTAGCAACAC AAGAGAAAGT TAATTATCA ATTCGTGGTA    | 3360 |
|    | ACGTTAAAGA GGTAGTAGCA CGTGGTGCAC ATCCATGTAT TATTTCTATG GAGGGTCTTG   | 3420 |
| 15 | AAAAAGAAGG CGACACTTAT GTCATTCCCTC ATGTACATGA ATTGTTAAGC CCAATTAGTAT | 3480 |
|    | CAGTGGTTCG ATTACAATTA ATTTCTACT ATGCAGCATT ACACAGAGAT TTAGATGTTG    | 3540 |
|    | ATAAACACAG TAACCTTGCT AAATCAGTTA CTGTGGAATA ATTCACTTT TTAGAATCAA    | 3600 |
| 20 | TCATGTATTA AAATTAAAGT ATATGGCACC CTTTITAGATT AATCGACTAG AAGGGTGCTT  | 3660 |
|    | TTTTAGGTGC ACTTtGCTTT TACTTCATCT TAATTGGCA GAAATGCGTa AAAATGAAGT    | 3720 |
| 25 | GTTTtATTTA TTTAAATAGT CTGACAATTA AGGGTGTTAT GTTAATATGA TTTTATGAGA   | 3780 |
|    | AGTATGGAGT AGCAATAAAG GGGTGACCTC GCATGTTAAT TCAATTAGAT CAAATTGGGC   | 3840 |
|    | GAATGAAGCA AGGAAAAACA ATTTTAAAAA AGATTTCCTG GCAAATTGCT AAAGGTGATA   | 3900 |
| 30 | AATGGATATT ATATGGGTTG AATGGTGCTG GCAAGACAAC ACTTCTAAAT ATTTTAAATG   | 3960 |
|    | CGTATGAGCC TGCAACATCT GGAACGTGTA ACCTTTTCGG TAAAATGCCA GGCAAGGTAG   | 4020 |
|    | GGTATTCTGC AGAGACTGTA CGACAACATA TAGGTTTTGT ATCTCATAGT TTAAGTGAAG   | 4080 |
| 35 | AGTTTCAAGA GGTGAAAGA GTAATCGATG TGGTGATAAG CGGTGCTTTT AAATCAATTG    | 4140 |
|    | GTGTTATACA AGATAITGAT GATGAGATAC GTAATGAAGC ACATCAATTA CTTAAATTAG   | 4200 |
|    | TTGGAATGTC TGCTAAAGCG CAACAATATA TTGGTTATTT ATCTACCGGT GAAAAACAAC   | 4260 |
| 40 | GAGTGATGAT TGCACGAGCT TTAATGGGGC AACCCAGGT TTTAATTTTA GATGAGCCAG    | 4320 |
|    | CAGCTGTTTT AGACTTTATT GCACGAGAAT CGTTGTTAAG TATACTTGAC TCATTGTTCAG  | 4380 |
| 45 | ATTCATATCC AACGCTTGCG ATGATTTATG TGAAGCAGCTT TATTGAAGAA ATAAGTCTTA  | 4440 |
|    | ACTTTTCCAA AATTTTACTG CTAAAAGATG GCCAAAGTAT TCAACAAGGC GCTGTAGAAG   | 4500 |
|    | ACATATTAAAC TTCTGAAAAC ATGTCACGAT TTTTCCAGAA AAATGTAGCA GTTCAAGAGAT | 4560 |
| 50 | GGAATAATCG ATTTTCTATG GCAATGTTAG AGTAAATATT TTGCAAAATA TAAGTAATAA   | 4620 |
|    | TGACAAAATT TAATTAGAT AAAATGGACA GTGGAGGGCA ATATGGATAA CGTAAAAGCA    | 4680 |

55

EP 0 786 519 A2

AAAGATGTCA TTAATCAATT GAGAGAGAAA GGATATAAAG TATTTTTGGC AACAGGACGT 4800  
TCGCATTCTG AAAATACATC AACTTG 4826

5 (2) INFORMATION FOR SEQ ID NO: 276:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 4846 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 276:

|    |  |      |
|----|--|------|
|    | GAATAAAAAAG TAAAAATFACT CGCCTTTGTT ACCTTTTACT TTATCAATA&A AATCAGTTGC | 60   |
|    | TTTTTCTTTT GCATTTTCAA CGAATTCCTT CGCTTTACCA GAAGCTTTAT CTCCTTTACC    | 120  |
| 20 | TTGCTTTTCT AAATTTTTAT TATCAGTAAC ATTACCTACT GTTTCCTTAA CATTACCTTT    | 180  |
|    | TGCTTGTCCA AATTTACTTT CGTCTGCCAT AATAAATGCC TCCTCGGAAT AATTAATATT    | 240  |
|    | TATATAATAT ACTTACTCCC TGA&AAAAATA TCTAAACACT TCACTTAAAT AATTTTGTAT   | 300  |
| 25 | ATTGATTTGA CGTCATTTTA TAACTAGCGA AATAGATTCA TCAATTAACCT GAGGGAAGTG   | 360  |
|    | GACTGAAATA ATAAAGAATC ACTAATGATT TATGATGTAT TAGTCACTAG CCATGTGAAA    | 420  |
|    | TTAAAAATAA GAATAAATGA GTAGCAGCGA TGCATATAGG ATTTTACTTT ATCCGTAATA    | 480  |
| 30 | GCATCTCATT CCTAAATATC ATATAAATAC CTGTTTAAAT TAAAAAGCCC AGCAACATCA    | 540  |
|    | CGTTACTGAG CCATTAATAT GATTTATTTA GCAGGAATAA CTGCACCATT GTATTTTTCA    | 600  |
|    | TTAATGAAGT CTGGAATATC TTTAGATTGT AATACTTCAA TTAATGCTTT GATTTTCTTA    | 660  |
| 35 | TCATCTTGAT GTCTTCTCTT AACAGCAATT AAGTTTGCAT AAGGATTATC TTTCCGCACT    | 720  |
|    | TCTACGACAA TAGAATCTTT TTTAGGATTT AGTTTTTGTG CTGATGCAAA GTTCCGAATA    | 780  |
| 40 | ATGATAACAG CATCAGCGTC TTCAATTTGA TAAATTTTAG GTAAGAATTC TGCTGATTGT    | 840  |
|    | TTATTATTAA ACTTAATATC TTTTFTATTC TCTGTAATAT CACTAAACCT AGCATCTTCA    | 900  |
|    | ATTTTTACGC CTTTTTGTAT TTTAATTA&A CCTGCATCAA CGAAGAATTT TAAGAAACGT    | 960  |
| 45 | CCTTGTTCAG CTGGATTATT AGACACATAG ACTGTGTGAC CTTTGGTAA TCTCTTTAAA     | 1020 |
|    | CTTTTACTAT TTTTAGAGTA TACAGCCATA GGTTCCTAAGT GAACATCACC GGCACCTTACG  | 1080 |
|    | ATTTTGTAA& CTTTATCCTT TTTCTCTGTG TTTAAATATG GTGTATGTTG GAAATAGTTT    | 1140 |
| 50 | CGCTCAATTT CACCTTTGTC TAGTA&AATTA TTAGGTGTAG TGTAACTGTT AATTTGTTTA   | 1200 |
|    | ATATCTAGTT CATA&ACTT TTTCTCAAT TCTGGTTTTC TTTCTTCTAA AATTTGACGA      | 1260 |

|    |             |             |             |             |             |             |      |
|----|-------------|-------------|-------------|-------------|-------------|-------------|------|
|    | TTACGCGAAG  | CTGCTAATAC  | AACCTGCAAAT | GTTAATACTA  | AAATAAGACC  | AAATAATTTT  | 1380 |
|    | TTCATAAAT   | GAAACCCCCA  | ATTATCTGTT  | TATCAAGTTT  | AITTTGTAAGC | CAATCCCCAA  | 1440 |
| 5  | TGAATTGGAT  | TATAAATACA  | ATAATTAAAA  | TAAAACTGTT  | TGATACTAAA  | ATGACATCAT  | 1500 |
|    | TTTGATTTCG  | AGTGAACCTT  | GTTAAGTATG  | CTAAATTTC   | TAAACCACGC  | GCACCAATTA  | 1560 |
|    | CACCTGCAAC  | TGCTGTTGAA  | CCAACATAAG  | CGATTGCTGT  | AACTGTAATG  | CCAGACACTA  | 1620 |
| 10 | GCCTGCGCAT  | AGCTTCAGGT  | AAAAGGACTT  | TACGAATTAC  | TGTCGAAGTA  | TTAGCGCCCC  | 1680 |
|    | TTGACCAAGC  | CGCTTCGATG  | ACACCTTTAT  | CAATTCTCTT  | AAAAGCAATT  | TCTAGGAGCC  | 1740 |
|    | TTGCATAAAA  | CGGTGtGtG   | CCAAATGCTA  | AGGCTGATGTA | CGCACGTCAT  | GGACCACTTA  | 1800 |
| 15 | TCGTTCCAAG  | TATCAAACTT  | GTTAAATGGAA | TTAATAATAA  | AATTAAAAAT  | ATAAATGGTA  | 1860 |
|    | TCGCTCTAAA  | TAAGTTAAACA | ATGAAGAA    | CGATAGAATA  | AAATAACCTT  | GCACCGATAG  | 1920 |
| 20 | ACTTACCTTT  | AGCAGACAAG  | ATAATAACA   | CACCTAAAA   | AAGACCAAGT  | ATAAATGCAA  | 1980 |
|    | ATATAGTTGA  | GACGACTGTC  | ATGTATAGTG  | TTTCGACTAT  | TGCAGTCCAA  | ACTTCTGGCC  | 2040 |
|    | ACTGAATATT  | AGGCATTGTA  | ATCATTTTCAT | TTATAATTTC  | ACTAAATGAT  | TTACCCATGT  | 2100 |
| 25 | CTTAACACCT  | CCATTTTAAC  | TGTGCGCTCA  | ATTAACCTCT  | TTTCGAATTT  | TCCGAAATCT  | 2160 |
|    | ACACTTGAAA  | TATATGGAAT  | ATGCAAGAACT | AAAAAGCCGA  | CTGTTCCATT  | TTTTGTATTT  | 2220 |
|    | TTAATATTGG  | CTTCTAAAT   | ATTAATTTTA  | ATATCATAGC  | CAGTTGATAG  | ACTCGATACA  | 2280 |
| 30 | ATAGGCTCGG  | TGTTGTGTGA  | ACACAGCGAA  | ACTAATCTAA  | CGATATAATG  | ATCTTTTCTT  | 2340 |
|    | AATGGCTCTA  | ATTCTGTTTAA | AGATGTTTGC  | AAGATCATCAT | TTAATTCGTC  | TTTCAACAAT  | 2400 |
|    | CGTTTGTGCA  | CAGTGTGTTG  | CGGATTTTCA  | AAAACCTGTG  | TCACCGGTCC  | TTTCTCTATC  | 2460 |
| 35 | ACTTTAACCAC | TTTCCATAAC  | TGCAACTTCA  | TCACAATATC  | GACGAAATGC  | ATGCATTTCG  | 2520 |
|    | TGCGTAATTA  | GTACAATTGT  | TAAATTTTGT  | TGTTCTCTAA  | TTTTTAGTAG  | TGATCTATAA  | 2580 |
|    | ATTTCTATCTG | TTGTTTGGCG  | ATCAAGTGCA  | CTTTTGGCCT  | CATCACAAG   | CAAGACCGTT  | 2640 |
| 40 | GGATCATTA   | TAACGCTGCT  | GCAATCCCCA  | CAGTTTGCTT  | TTGTCCACCT  | GATAACTCTG  | 2700 |
|    | ATGGATAAGC  | CTTTTCTCTA  | CTTTTAAAC   | CGACGAGTTC  | GACAAGTTCT  | AATGCTTTTT  | 2760 |
| 45 | GCTTAGCTCT  | CTTCTAGGG   | ACACCTGCAA  | TTTCAACGGG  | AAACATAATA  | TTTTTTAAACA | 2820 |
|    | CAGTCTTGTA  | CCATAACAAA  | TTAAATATGTT | GGAAGATCAT  | ACTTACTTTT  | TGTTTTTTTG  | 2880 |
|    | CTCTTAATGCT | ATTTTGGAC   | AATTGACCTA  | TATGGTCTCC  | ATCTATAATA  | ACTTCCACCTG | 2940 |
| 50 | TTCTAGCGCG  | TTCTAAATGA  | TTAAACATTC  | GAACTCAAGT  | ACTTTTCTCT  | GCTCCAGAAA  | 3000 |
|    | AACCAATGAC  | GCCATAAATC  | GATCTGCTCT  | GAATCGATAA  | ATTAACGTGA  | TCTACGACAA  | 3060 |



## EP 0 786 519 A2

|    |   |      |
|----|---|------|
|    | TTCTCTCCTG TGTGTCTTAA TAAAAATAAA AATGCTTTCT CAATATCGAT AGAAAAATTG   | 3180 |
|    | AGAAAGCAAT AGTAGTATTG TTTCTCTCAT CTTCAAAAGT TAAAACTTTA TGTGAATTGG   | 3240 |
| 5  | CACCATTTCT ATATAAGACG GTTGCGGGGC TTCGTAGGGC ACATCCCTCC ACCACTCTCG   | 3300 |
|    | ATAAGAGTTT ACGCATCATT TAATTTGTAT TAATCCTAAC ACCTTAGTAA AATTTTCGTCA  | 3360 |
|    | ATAACTATTT TAAATTTTCT AACAAATCAG TCACCGATTT AATGCATAA ATTCTGTTTTA   | 3420 |
| 10 | CTTCTTTATC TTTATTCATC AACAAATAAA TCGGCGTAGA CATGATTTCG ATATCTTTAC   | 3480 |
|    | AAAACCTGAGG ATAAAAGTTT AAATCTATTT TCAATATGGT TAACGTCAAT ATTTCAATTAG | 3540 |
|    | CAATGTCTAA CATTCTTTCT GAAACCTTAC AAGTACCACA CGTTGGTGTA TAACCAAGA    | 3600 |
| 15 | TTAAATGTTT GTCTTCTCA TAAATGTAG TTACATCTTT GATGCTAAT GAATTATTCA      | 3660 |
|    | TTTACTAAAA CTAACCTTTC ATTATTTATA TTGGTAAAA GAGGTGTTTC TTTCTTACAA    | 3720 |
|    | GTAAGCCAT GTTTTGAAAG TACATGCGCC AAATATTGTT TGGGGCAATT CGCAACTTGA    | 3780 |
| 20 | CAGTAAGTTT TATCAATAAA TATATGTTCA CTTCACCTCA AATAACGTTT AAACCAATTT   | 3840 |
|    | CTAATTCGAT CTCTTCGTG ATCAGAATCG GCTAATACAA AAACCTGTTT ATCATAACGT    | 3900 |
| 25 | GATTCTATCA TATCATCAAG CTTATCTATA CTCATTGTTT CATGAGTACA AATAATATTG   | 3960 |
|    | ACTGGTCTG CAATAACCTG TTGCAACCTT TTTTATCAG ATTTTCCTTC AACAAATTATC    | 4020 |
|    | ACTTTATTTA CAATAGCCAT CATCATCACC CTTTAAAAAT AATAACATC TGTCACTGTA    | 4080 |
| 30 | TCATTTGAGG AAATTTGGTAT GAATAAAACA TAAATCAACA AAAATTTAAA ATTTCAATTA  | 4140 |
|    | ATAATAATTA CAAACTCAAT GTTTGACTAG CTGGAACAT TAAACATAGC AGACAAAGGC    | 4200 |
|    | TAAGTCAAAA ATCAACATCC TAAAATCTAC AATGTTATAT TAACATAGT TAACCAAAAG    | 4260 |
| 35 | AAAATACACC TATAACAAAC TTTTCAATTA TAGCGGGGCC CCAACACAGA AGCTGATGGT   | 4320 |
|    | AAGTGAAGCTT ACAATAATGT GCAAGTTGGC GGGGCCCCAA CATAAGAAA TACTTTTTCT   | 4380 |
|    | TTAGAAATTA GTATTTCTTA TGCATGAGTT TTAATCATGT ATTCCTATTT TTAAATACAC   | 4440 |
| 40 | ATTAGCTGTG GCTTATGAAA ACAGGCTGGG ACATAAATCA ATGTTCTATG CTCTACGAAG   | 4500 |
|    | TTATATTGGC AGTAGTTGAC TGAACGAAAA TGCGCTTGTA ACAAGCTTTT TTCAATTCTA   | 4560 |
| 45 | GTCAGGGGCC CCAACACAGA GAATTTTCGAA AAGAAATTC TACAGGCAATG CAAGTTGGGG  | 4620 |
|    | ATGGGCCCCA ACAAGAGGAA ATTGGATTCC CAATTTCTAC AGACAATGCA AGTTGGGGTG   | 4680 |
|    | GGACGACGAA ATAAATTTTG CGAAATATC ATTTATGTCC CACTCCCTAG ATTGATCTAT    | 4740 |
| 50 | AGATACTACA CTTATTAAG TAATATATTT TTATGATICT CTAGCTGCA ATCCCATGAA     | 4800 |
|    | TACATGTAAT CATCAAACCT CATAGCCTCA AGGTCAAGT ATTCA                    | 4846 |
| 55 |   |      |

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1843 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## 10

|    |             |             |             |             |            |             |      |
|----|-------------|-------------|-------------|-------------|------------|-------------|------|
|    | AACAAGACA   | CATTCGAACA  | TGAACCATCA  | GTAAGAAGCTG | AAGATATATC | AAAAAAGGAG  | 60   |
|    | GATACACCAA  | AAGAAGTAGC  | TGATGTTGCT  | GAAGTTTCAGC | CGAATTCGTC | AGTCACTCAT  | 120  |
| 15 | AACGCAGAGA  | CACCTAAGGT  | TAGAAAAAGCT | CGTTCGTGTT  | ATGAAGGCTC | TTTTGATATT  | 180  |
|    | ACAAGAGATT  | CTAAAAATGT  | AGTTGAATCT  | ACCCCAATTA  | CAATTCAAGG | TAAAGAACAT  | 240  |
|    | TTTGAAGGTT  | ACGGAAGTGT  | TGATATACAA  | AAAAAACCAA  | CAGATTTAGG | GGTATCAGAG  | 300  |
| 20 | GTAACCAGGT  | TTAATGTTGG  | TAATGAAAGT  | AATGGTTTGA  | TAGGAGCTTT | ACAATTAATA  | 360  |
|    | AATAAAATAG  | ATTTTAGTAA  | GGATTTCAAT  | TTTAAAGTTA  | GAGTGGCAAA | TAACCATCAA  | 420  |
|    | TCAATATACCA | CAGGTGCTGA  | TGGTTGGGGG  | TTCTATTATTA | GTAAGGAAAG | TGCAGAAGAA  | 480  |
| 25 | TTAATTAACTA | ATGGTGGAAAT | CCTTGGGGAT  | AAAGGCTCTG  | TAAATTCAGG | CGCATTTAAA  | 540  |
|    | ATTGATATCTG | GATACATTTA  | TACAAGTTTC  | ATGGACAATA  | CTGAAGACAA | AGCTGGACAA  | 600  |
| 30 | GGTTATAGAG  | GATACGGAGC  | TTTTGTGAAA  | AATGACAGTT  | CTGGTAATTC | ACAAATGGTT  | 660  |
|    | GGAGAAAATA  | TTGATAAATC  | AAAAACTAAT  | TTTTTAAACT  | ATCGGCACAA | TTCAACTAAT  | 720  |
|    | ACATCAGATG  | GAAAGTTTCA  | TGGGCAACGT  | TTAAATGATG  | TCATCTTAAC | TTATGTTGCT  | 780  |
| 35 | TCAACTGGTA  | AAATGAGAGC  | AGAATATGCT  | GGTAAACCTT  | GGGAGACTTC | AATAACGAT   | 840  |
|    | TTAGGTTTAT  | CTAAAAATCA  | GGCATATAAT  | TTCTTAATTA  | CATCTAGTCA | AAGATGGGGC  | 900  |
|    | CTTATTCAGG  | GGATAAATGC  | AAATGGCTGG  | ATGAGAGACTG | ACTTGAAAGG | TTACAGAGTT  | 960  |
| 40 | ACTTTTACAC  | CAGAAGCGCC  | AAAAACAATA  | ACAGAAATTAG | AAAAAAAAGT | TGAGAGAGATT | 1020 |
|    | CCATTCAAGA  | AAGAACGTAA  | ATTTAATCCG  | GATTTAGCAC  | CAGGGACAGA | AAAAAGTACA  | 1080 |
|    | AGAGAAGGAC  | AAAAAGGTGA  | GAGACAATA   | ACGACFCCAA  | CATCTAAAAA | TCCATTAACT  | 1140 |
| 45 | GGAGWAATTA  | TTAGTAAAGG  | TGAAYCGAAA  | GAGAGAAATCA | CAAAAGATCC | GATTAATGAA  | 1200 |
|    | TTAACAGAAT  | ACGGACCAGCA | AACGATAACA  | CCAGGTATCC  | GAGACGAATT | TGATCCGAAG  | 1260 |
|    | TTACCAACAG  | GAGAGAAGA   | GGAAGTTCCA  | GGTAAACCAT  | GAATTAAAGA | TTCAAGAACA  | 1320 |
| 50 | GGAGAYGTAG  | TTAGACCACC  | GGTCGATAGA  | TGACAATAAT  | ATGACACTGT | AACGAGAGAC  | 1380 |
|    | TCGAGTTGTAG | AAAAAGAGA   | ATTCTCATTC  | TAGAAGAAGAC | GTAATTTTAA | TCCTGATTTA  | 1440 |

CCAACACTAA AAAATCCATT AACTGGAGAA ATTATTAGTA AAGGTGAATC GAAAGAAGAA 1560  
 ATCAACAAAAG ATCCGATTAA TGAATTAACA GAATACGGAC CAGAAACGAT AACACCCAGGT 1620  
 5 CATCGAGACG AATTGTGATC GAAGTTACCA ACAGGAGAGA AAGAGGAAGT TCCAGGTAAA 1680  
 CCAGGAATTa AGAATCCAGA AACAGGAGAT GTAGTTAGAC CACCGGTCGA TaGCGTAACA 1740  
 AAATATGGGa CCGTGAAGA GAGACTCGAT TgTAGGAAAA AGaGAATTc CaTTCaGGAA 1800  
 10 AGACGTaATT TaTCTGTTT AGCACCCGGG GCAGAAAAGT TAC 1843

(2) INFORMATION FOR SEQ ID NO: 278:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 8536 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 278:

TAAACAGCGC GTGTACTTGT GATTCCCCCT TCTTCTATTT TACCCACCCG GGAATAATA 60  
 25 CTTTTCGCGA TTCTTACTT GaACAAGCAA TATTTTATCA GCTGTTTCTT CAACTAAACA 120  
 GACACATTTA ATCATCTTTG ACACCCCAAC TTTGTGAAAT CAATTTTTCa AATTATACTG 180  
 30 TACAATTATG TTATCATATA TGAGTAGTTA TAGCGCAAAA CGTTAGCAAT TCAGCGCACC 240  
 CAACTTTTCA TATAAACAGA AGATACTAGG GGAATTATIT ATTATGGCAa AACgTTCCAA 300  
 ATCACAACTG TTATCAAGTT TACTAAATGT CGCAGGTTTC ATAGTCGACG GCTACAATGG 360  
 35 CTATAAATAT CATGCTAAAA ATAAAAAAT AGTATATCTT TCATTAGGIT TAAGCACTGT 420  
 AGGAACCGTG TTAGACTTTT ACATTTCAAT TAAGTCACCA CGTAAGTTCA AAAAAAGCAGT 480  
 GGCAGTTGTT ACTTTAATAA CAAACGGTGC TAGATTATTT ACAAGCATTc GCAAAGTAAA 540  
 40 ACATGAATAC TAATTCAGAA AAGGATTGGT CGAACATAGA ACATGAAGTT CATTCGACCA 600  
 ATCCTTTTTT ATATACAAAA ATTCTATTA CAACAATCAC GACTTTGATA GCCCCGCCAT 660  
 AAAATTTAGG ATTCAATCCA ACTTTTCAGC TTGTGAaATG TAATAGGaAT TCATTATATA 720  
 45 TTTATATACG TAAGACTTTA GTGAATATAT CTATAATTAT TTACTTGGTA AGCTGGTACC 780  
 GTTCTGTaAG TTAATAATCC TGGTGCAGCT GAATAGTTCA TTCTGAAAC TAAATACTA 840  
 50 CCATCGTTAT TTACAGTTTC TACAAACATA ACATGACCAT AGTAACCTAC ATCTGTTTGA 900  
 GCGATAGAaC CTACAGTAGG TCTATTGTCG ATAGTGTAAC CATCTGCTGC CGCTGCGTTA 960  
 TCCCAGTTAT TAGCATTCCA CCAATAAGTA CTAATACCTT TACCAATTTc AGCACGACGA 1020

EP 0 786 519 A2

|    |             |            |            |            |             |             |      |
|----|-------------|------------|------------|------------|-------------|-------------|------|
|    | GTATTGTAAC  | CTCTATTGTG | CGTTGTTGCA | GATCCTGAGT | TGCTAGATGC  | aTTACCAGTT  | 1140 |
|    | ACTTTCAATT  | TTTGACCCGG | ATATATAAAG | AAATTATTTA | AGCCATTTAA  | GCTCATAATT  | 1200 |
| 5  | TTTTGATAAG  | TTGTACCATA | TTTTGATGCA | ATTAATGACA | ATGAGTCACC  | TGCTTGTACT  | 1260 |
|    | GTATAGTATG  | ATCCGCCACC | TGAGTTCGTT | GATGGACGGC | TACTATTGCT  | CGCAGCGTTA  | 1320 |
|    | CTTGAGCTAG  | CAGTACCTGA | TACTTTTAAT | TTTTGACCTG | GATAAATAAA  | GAAATTATTT  | 1380 |
| 10 | AAACCATTAA  | GTGCATAAAT | GTTTGGTAA  | TTGTACCAT  | ATTTTGATGC  | GATTAATGAT  | 1440 |
|    | AATGAGTCGC  | CTGCTGTAC  | TGTGTAGTAT | GATCCGCCAC | CTGAGTTCGT  | TGATGGACGG  | 1500 |
|    | CTACTATTAC  | TGCTAGAATT | ACTTGAGCCA | GATACTTTTA | GTACTTGGTT  | TGGGAAAAAT  | 1560 |
| 15 | AGATTAGATG  | TTAAATTGTT | TAATGACTTT | AATTTAGCAA | TCGAAATCCC  | ATACTTATTT  | 1620 |
|    | GAAATTGCCC  | ACACTGATT  | ACCCGGTTTT | ACTGTGTGAG | TTGTAGCCGC  | ATTTGCTTGA  | 1680 |
| 20 | GTTCGCCCAA  | CAGCGCTAAT | CGCGCTTGTC | CCAATAATAG | CTGCAATTAC  | TTTTTTTTGC  | 1740 |
|    | ACTTTAAAAAT | CCTCCTCTTG | CTTAACCTTC | CTAACATTCT | TTTATCCGAA  | TTATGAATA   | 1800 |
|    | CTACATCATT  | ATACGATTTT | ATTATGTATA | ATAGGTTGAT | GTTTGATGAC  | ATTATGaTTA  | 1860 |
| 25 | CAAAAAAATC  | ATATACTGTA | TCATCAAATT | TTATAATTAT | CCCTTAAAAT  | TATTACAAC   | 1920 |
|    | TATTAGATT   | TACAATATCT | AAATTATTAC | AATTTCAATA | TATTTCACTA  | TAAAAATGATT | 1980 |
|    | ACAATCCCTT  | TCTCTATTGG | AAATAATTTT | ATTCTCCAAC | AATAACGCCC  | TACAACACATA | 2040 |
| 30 | AGCATGAAC   | TTGCTGTAG  | GACGCAATAT | AATTTATTTT | GCTAATGACG  | TTTCTATTGC  | 2100 |
|    | CTTAACTCTCA | TCTTTAGATA | AATTAACAGG | TTTCTCTCCA | TCTTTGACAT  | CTTCTTGCAA  | 2160 |
|    | CGCTTTTGA   | GCTTCTTTTG | AATGATACAA | TTCTACGATT | TTAGCATATG  | TTTTGTTATC  | 2220 |
| 35 | CAAGTCTTIG  | TCATTAAC   | CAACAATATT | AATATATGGC | TTTACTGCAT  | CTGAATTGTA  | 2280 |
|    | TTTTTCTAAA  | AATATCGGAT | CATTTTTAGG | ATCTTTACCC | GCTTTAGTTG  | CTACACCGTT  | 2340 |
| 40 | ATTAATAACT  | GCAATATCGA | CATCAGATAA | AGCACGTGCA | GTTTGTGTG   | CATCTACTGC  | 2400 |
|    | AGTAATTTTT  | AAATGTTTTG | GATTTGACGT | TATATCTTTC | ACCGTGCCCTG | CTAATCCGAA  | 2460 |
|    | ATCTTTTTTC  | AGTTTTATTA | AACCAGCTGC | TTCTAATAGT | TTAAGTGCAC  | GTGCTGGT    | 2520 |
| 45 | TGACACATCA  | TTTGGAATGA | CAACTTTAGC | ACCATCTTGA | ACCCTTTTGA  | CATCTTTAAT  | 2580 |
|    | TTTATCTGAG  | TAAATGCCCA | ACGGTGCTAA | AACGTGTGTA | CTTAATGCTG  | AAATCTTTGT  | 2640 |
| 50 | TCCTTTATGC  | GCCTTTTAT  | ATTGATCTAA | AAATGCAAAA | TGTTGGAATG  | CATTATATC   | 2700 |
|    | AATATCACCA  | TCATTTAATG | CTTTATTGCG | TAAATGTAA  | TCTGAGAAGT  | GCTTAATCTC  | 2760 |
|    | CACATCAATA  | TCATCTTTTT | TAGCTAATTC | TTTAACCTTC | TCCCAAGCCT  | TAGTGTCAAT  | 2820 |

55



|    |  |      |
|----|--|------|
|    | TAAGAGCATT ATATGTAAAA TTGCATATAT CGTCAATACA ATTTGCCGAA TTTTCTAAAA  | 4740 |
|    | AATTAAAAAA TAAGTAATTC ATGTGACAAAT GACGAATGTG GAGACTACTA TGACATTTAT | 4800 |
| 5  | CAAAATTAAT CCATAAAAA GTCCACCAAT CCTCCACAAC GCAATTACTA AATATTAAACA  | 4860 |
|    | TGCACAAAA AAGCACTAGC ATATTCAAGA ACAACAAAG TTGAACCTAA AATATATGCC    | 4920 |
|    | AGTGTGCTA TTAATTATAA AGTATCTAGT GCTTGTTTTA AATCATCGAC TAAATCTTCA   | 4980 |
| 10 | GTATCTTCAA TACCTACAGA AATTCTTACA AGTCGCTCG TAATACCTTC TTTAGCTCGA   | 5040 |
|    | ATATCTGCTG GAATGGATGC ATGTGTCATC AATGCAGGTA CTGAAATTAA ACITTCCACT  | 5100 |
| 15 | GCACCTAAAC TTTCAGCTAA TGTGTAATAC GATGTTGCTT TAATCAATTG TTTGGCACTT  | 5160 |
|    | TCTGTATTTT TCACCTCAA TGCAATCACA CCTGTATGGC CATCGCTTGC AGCCATATGG   | 5220 |
|    | ACATCATGAT TTAATGACTT TTCAATACTT GGATGGAACA CTTGTTGCAC AGCTGGATGT  | 5280 |
| 20 | GCTTGTAACA TTTTAATAAT TTCAATAAGC CTGCGATTAA TTTGTTCCAT ACGTAAACCT  | 5340 |
|    | AATGTTTAA TACCCCTCAC AAGTAAATAG CTATCTTGAG GTCCTAAAAA GCCACCTGTT   | 5400 |
|    | GAATTTGAAA TAAATGCTAA ACGTTCTGCA AGCTTGTGAT CCGATGTTGC AACTAAACCA  | 5460 |
| 25 | GCAACGACAT CACTATGTCC ACCTAAATAT TTCGTTGCAG AATGTAAGAC AATATCGATA  | 5520 |
|    | CCTAAATCTA ATGGATTCTG ATAATAAGGT GTCATAAATG TGTATCAAC AACTGAAATC   | 5580 |
| 30 | AAACCGTGT CTTTCGCAAT TTCAGCAGAC TTTTAAATGT CAGTAACACG TAATAATGGA   | 5640 |
|    | TTAGAAGGTG TTTCATAAAA CAACATCTTT GTTGTGGGCG GTATCGCTTG TACAATTGAA  | 5700 |
|    | TCGTATGCG TGTATCTAC AAAATCCACT TCAATGCCAA ATCGTGTAAT TACTTTTGTG    | 5760 |
|    | AATGCGGAT AAGTACCGCC GTATACATCT GAATTTAAAA TAATATGATC TCCTTTGTCC   | 5820 |
| 35 | AACAGCATAA CAACTGCACT GATTGCTGCA ACACCTGAAC TAAATGCAAA GCCATGTTTG  | 5880 |
|    | CCATTTTCTA ATGTCGCAAT AACGCTTTCT ACAGAACTTC TTGTTGGATT CGCAGTACGA  | 5940 |
| 40 | GAATATTCAT ATCCTTGACG TAAATCACCA ATATCATCTT GTAAATATGT ACTTGTTTGA  | 6000 |
|    | TAAATTGGTG TTGTAACGGC ACCTGTATAA TCGTCTGTTG TGTGCCACCC ATGAATTAAT  | 6060 |
|    | TTAGTTTCTT TGTTCAATTAT TATTCTCCTC ATAATTAAT ATTTGCTTAG ACATATATCG  | 6120 |
| 45 | ATCACTACCA TCTGAAATA CGACAACAAT CGTACCTTCA GATAATTGCG CTTTAAAT     | 6180 |
|    | CAATGCACCT TGTAAATGCT CACCTGAGA ACTGCCTACT AACAAACCTT CATTTATAGC   | 6240 |
|    | CAAACTTTTG ACATTTCGAA AGGCATCTTG ATCTTTAATC GTAAATATCC CATCTACAAG  | 6300 |
| 50 | ACGTCTCTCT AAAAATATCG GCCATTCTCT AGAACCGATA CCTTCAGTGT CATGTGCATG  | 6360 |
|    | AGCTGGCCCT CCATTTAACA CGGACCCCTC TGGCTCAAAG GCATAACATT GCACGTGATG  | 6420 |

EP 0 786 519 A2

|    |             |             |             |             |             |             |      |
|----|-------------|-------------|-------------|-------------|-------------|-------------|------|
|    | ATAATCAATT  | TGCTGTAATG  | CTGAAGTCAA  | TTCCGGTCCC  | AATGTATGAA  | AATATGTATC  | 6540 |
|    | CGGATTATGT  | TCGGATTCAA  | ATTGATTTCAT | ATAAACGGCA  | CCATATTTTT  | CAGCATAGGA  | 6600 |
| 5  | ACGTGCGAGCT | AATTGTGCCC  | CATGCATACC  | TTCAGACTGA  | CTCGTCCTTG  | AAACTTCTGC  | 6660 |
|    | ACCAAGCGCT  | ATCATAATAT  | TAATCTTTTC  | TTCTGAAAAA  | CCATACGGCG  | CAAAGATCTT  | 6720 |
|    | ACATTTCAAA  | TGATGCTTAT  | TCGCTGCAAT  | AGCTAACCCCT | ATGCGTGTAT  | TACCAGCAGT  | 6780 |
| 10 | CGCTTCAACA  | ATAGTTTGAC  | CTGCACGCAC  | ACGCCCTTCT  | TGAATTGCCT  | TCTTACTTAA  | 6840 |
|    | ATATTTCCCG  | AGTCTGTCTT  | TAACTACTGCC | TCCAGGATTG  | CATTGTTCAA  | GCTTGGCATA  | 6900 |
|    | AATTTTAACT  | TTATCATCAC  | TATAATGTTT  | TAACTAGTCT  | AATGGTGTAT  | TGCCAATTAA  | 6960 |
| 15 | ATCATAAGTA  | ATCATAGATG  | CACCCCTCATC | TGACATGCCG  | ATCAAAATGAA | TGAAACCTTT  | 7020 |
|    | CTTCATGTCT  | CAATTTTAAAT | TCTTACTTTT  | CAGATAAGAA  | TTATAAACGA  | CATTTTGTTA  | 7080 |
|    | TTTTGCAATT  | ATCTAAGTTT  | CGATTAAATC  | AGAACCAGTA  | CTAAATTTTC  | AATTCCAAAC  | 7140 |
| 20 | AAAAAACAC   | CTGAGCAACA  | CAAAATACCTG | TGTGTCAAGT  | GCTTCTATAT  | ATTAATCTAAA | 7200 |
|    | TAATTGCACG  | ATAAAGACTA  | AAATAATAAC  | GACAGGCATC  | GCATACTTAA  | TTAAGTAATA  | 7260 |
| 25 | CCAACCACTG  | AATAATCTAA  | ATCGATCTTT  | ACCAAAATAT  | TGTTGTAATA  | ATTTTTTATC  | 7320 |
|    | TAATAATTGT  | CCTAGACAAA  | CGGTAGTACC  | TAATGCGCCT  | AATGGCATCA  | ATACATTGCA  | 7380 |
|    | AACGATGAAA  | TCCATATTAT  | CAAAAATCGT  | TCCCGCACCG  | AATCTTACAT  | CTTTTAAAGAT | 7440 |
| 30 | ACCAAAAGAT  | AAGGTTGCTG  | GAATACTAAT  | GATAAATACT  | AAAATACTAC  | CGATCACTGC  | 7500 |
|    | GACTTTTTTA  | CGTTTTGTAT  | TGTCATTCTT  | CGTGAAGTTA  | GAAACATTTA  | ATTCTAATAA  | 7560 |
|    | AGAAATAGAT  | GACGTTAAAG  | CCGCAAAATA  | GAAACAGCACT | AAGAATCCCA  | AATAGAATAA  | 7620 |
| 35 | TGTGCCTAGA  | TGCATTTGAC  | TAAAGACCAT  | TGGCAGTACT  | TTAAATAATA  | ATCCAGGCCC  | 7680 |
|    | TTCTTGTGGT  | TCATAGCCAA  | AACTATGTAA  | AGCCGGAAT   | ATAGCTAGAC  | CTGCCAATAC  | 7740 |
| 40 | AGATACAAAG  | ATATTCATAA  | CAACGATAGA  | AATAGCTGAT  | GACTTAATCG  | TCATGTCCTT  | 7800 |
|    | AGAGGCATAA  | CTGCATAAG   | TAATCATACC  | TGTAGTTCCT  | AATGATAACG  | TAAAGATGA   | 7860 |
|    | TTGACCTAGC  | GCAAAACAAG  | TGCCATCAGC  | AGTAATCTCT  | GATACTCTTG  | GTTGTAAAAAT | 7920 |
| 45 | AAATTTCAAC  | CCTTCTAAGA  | CGCCATCTAA  | TGTTAAAGAC  | TTAATCACAA  | TGACGATTAA  | 7980 |
|    | AAAGACAAAC  | AGCAATGGCA  | TCTAATCTTT  | CGATGCCCTT  | TCTAATCCTT  | TTTCAACACC  | 8040 |
|    | TAAACATGACA | ATAATCATCG  | TAGCGAATAT  | GAATATACCT  | TGCCCTAGAA  | CGGTTAACCA  | 8100 |
| 50 | AGGATTTGAT  | ATTACCGCTT  | CAAAATTCAT  | TTCTTGGAGA  | TGATTGATGC  | GTTGAAATAT  | 8160 |
|    | AACATAATTGC | CATAATACTT  | GTCCGATGTA  | AATGACAATC  | CAACCACCGA  | TAACTACTATA | 8220 |

55

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 4328 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

|    |            |            |            |            |            |            |      |
|----|------------|------------|------------|------------|------------|------------|------|
|    | GCTTTGGCCA | TTTATGTGG  | CGATTGAGAC | AATCGTGKT  | TGCTCTATTT | GATGTGTGAT | 60   |
| 28 | TTCAACTGGT | AATTCAGTT  | GCGATTGAAA | TAATGGCAAC | TTTTCCCAAT | CATTAAACAA | 120  |
|    | TAATTCAATA | CCTGCTATGT | CTAATACTTT | AGCACGTGCA | TCATCAACAA | GACGCGCTTC | 180  |
|    | CAATTGATTT | GCTTCTTCTT | TAATACCTGG | TGACGTACTT | TCTAATATCA | AATTAGATAT | 240  |
| 30 | AGGGATGTGA | CCATTAATTG | CATAATATAA | TGCAACACGC | CCACCATTG  | AATATCCAAA | 300  |
|    | CAATGTTATT | GATTTATCTT | TATATTTCAT | TAAAAATCGG | TCTAACACAG | TCGTAATATA | 360  |
|    | ATCAAAATTC | CACGTTTCAT | CCATTGAAGA | CTGATCTTCG | CCATGGCGCT | GTAAGTCTAT | 420  |
| 35 | AGTGATGACA | TGATAGTTAT | CAGTAAATTT | TTCGATGTGA | TTATGATAAG | TACGGCTGTC | 480  |
|    | GCTAAGAAAT | CCATGCAGAA | ATACTAAAAA | TTGATTGGTC | TCAACGTTTG | CTTCATAAAA | 540  |
|    | TTTATAATGT | GTGATGAATC | ATTTCACTCA | ATTTCTGGTA | TAAAAATTGA | TGCTGTTTAA | 600  |
| 40 | AGTTATCTTC | GCGATTGGTT | ATCAATTCAT | AAATCGTGCA | AGTTTCAGAT | AACAATGTGG | 660  |
|    | CATTTTAAAT | TTCTGAACAA | CTGTTAAAAA | GTTTAAAAAT | GAATTGATAT | AACCTAGCTG | 720  |
|    | TATACTCGAA | ATCCAATCCC | GTCGGGTGTC | CAACAACCGC | TTCAAAATAG | TCAGTTGCAC | 780  |
| 45 | TTTCTTTTGG | TGCTAAATAT | GAAAAAATAC | CGCCACCATC | GTTGTTCAAT | ATAACAATAT | 840  |
|    | TCATCTGAAT | ATTATTTAAT | TTTGCAATTA | ATGATCCATG | CATATCATGA | TAAATGATA  | 900  |
|    | AATCACTTAT | CAATAATGTT | ATTGGTTTAT | GCACAGCCAT | ACCCAGTGCA | GTTGAAACGA | 960  |
| 50 | TACCATCAAT | ACCATTGCGA | CCACGATTGC | CATAGACATC | TATATTTTTA | TTCAATAACA | 1020 |
|    | AGTTATCTAC | ATCTCTGATA | GGCATACTAT | TACTAATAAA | TAATGCATCT | TTTTCAGATG | 1080 |



|    |            |             |            |            |             |             |      |
|----|------------|-------------|------------|------------|-------------|-------------|------|
|    | TAATTTCTTT | ACGCCCTTTT  | TTCTCTAAGC | ATTGCCATTT | TTCTAACCAA  | CTTACGGAT   | 1200 |
|    | TAACGTGCT  | GTCTTCCATT  | AATGACCTAA | AGAAATCAIT | CGCAGAAATC  | TCATATGAAA  | 1260 |
| 5  | TATCTGGCGC | TATCGGAAAG  | ACATCAATCT | TATCATTGTT | TTGCACTAAT  | ATTGTAAATG  | 1320 |
|    | CATCAGTTTT | CTTTAACCAT  | TGATTTAACT | TTTTAGAAAT | CACTGGTTTC  | CCAACACGAA  | 1380 |
|    | TTACGAAATC | CACATTTAAG  | TCTAAGCCGC | TTCTAAACAG | CAAAATCATAT | GTACAGATAA  | 1440 |
| 10 | CATTGCGATG | ATCAAAATTTT | CTTAAATGAC | TTAAAGGATC | AGCTAAAATA  | GGCAAAATCAT | 1500 |
|    | ATATGCTTGA | ATACGTTAGT  | ATTGATCAA  | CTTCTGGTG  | CTGCATATCC  | CCTACAATAA  | 1560 |
|    | TTAAACCTTT | TTTCTTATTT  | AAAATGTGTC | TTAATGCCGA | TGCATCTATA  | CTTTTTTGAT  | 1620 |
| 15 | AGTGGGTAA  | AATCTTCATC  | TCAGAAGTTA | ACAATTCGT  | TGCATTCAAA  | TCAGGTGTTA  | 1680 |
|    | ACGGATCTCT | AAATGGCAAG  | TTAAATGAA  | TTGGCCCTTT | ATGTGTTCCA  | TATAAATATT  | 1740 |
| 20 | GACTAGCAAT | TTGCATTGGA  | TAGTAAATTG | CATCAATGGT | CTCTTTACTA  | TCATCGCGAA  | 1800 |
|    | TAGGCATATC | GAACCTATAA  | CTTACATAAT | TATTAACAT  | ATTTACTTGA  | TTAATCGCTT  | 1860 |
|    | GTGGTGGCC  | TACACTTCTT  | AATTCATGCG | GACGGTCACT | TGTTAAAACG  | ATTAAAGGAA  | 1920 |
| 25 | TTCTACTAAT | TTGGCTTTCA  | GCAATTGCAG | GCGTATAATT | CGCTGCTGCT  | GTACCTGACG  | 1980 |
|    | TACATAATAT | AGCGACAGGT  | CTTCACTGCT | CTTTAATTAA | CCCAACTGCA  | AAAAACGCTG  | 2040 |
| 30 | CACCTCGCTC | ATCGGGGTGT  | ATCCATGTTT | TAATATTGG  | ATGTGCTTCA  | AATGCAGTG   | 2100 |
|    | CAAGTGGCGT | TGAGCGTGAT  | CCCGGACTGA | TAACACTTTC | CCTTACGCCG  | TACGCATATA  | 2160 |
|    | ACTCAGATGC | AAATGTAAAA  | ACTTGCTTCG | TTAAAGCTGC | TTTATGATTT  | CCCATTTCATA | 2220 |
| 35 | TCGACTCCTA | ATGCATTCAT  | CATAGGTGTG | AACCTAAGGT | TCGTTTCTGC  | CAATTCACATA | 2280 |
|    | TCTGGATCAG | AATCTTTAAC  | AATGCCACAC | CCAGCAAATA | AAGTGTCTTG  | TGCTTTCTTA  | 2340 |
|    | ATAAGCATCG | AACGAATTGC  | AACAATAAAT | TCACAATCAT | CGTATATATC  | TATATAGCCA  | 2400 |
| 40 | AACGGTGAC  | CATATAATCC  | TGCGTACCA  | AATTCTTTCT | GCTCAATAAA  | ATCCATTGCA  | 2460 |
|    | AATCTTTTG  | GATAGCCACC  | TAAAGCAGGT | GTTGGATGTA | AATTATCAAT  | TAAACTAATA  | 2520 |
|    | TACGAATCAT | CCTTCAGTGG  | CGCCTTTATT | TCAGTGTACA | AGTGATATAA  | ATGATCATTT  | 2580 |
| 45 | TTTAGAATTT | TAGGCGTCTT  | ATCATAATGT | AATTCAGTGA | TATAAGGTTT  | AATATCATGT  | 2640 |
|    | AAAATACTGT | CAACAACAAA  | TGATGTTTCG | ATTAAGTTTT | TATTATCTTT  | TAAAAATGCT  | 2700 |
|    | TCAACATTTT | TTGTATCTTC  | GTCTCATCT  | TGTGAACGTT | TAATTGTACC  | TGCTACAGCT  | 2760 |
| 50 | TTAGTCGATA | GTAATTTTATT | ATTGACCTTT | ATTAATGTGT | CAGGTGTTTG  | TGAAAAAGAA  | 2820 |
|    | ATAGAATCTT | GTGATTCATA  | CAAGAATATA | TAACGTGTTT | TTTCTTTAGA  | ATATGCTTGC  | 2880 |

EP 0 786 519 A2

ACAATTTTT CTTCATTATT AATAGATTCT ATAGCTTCTA CTACAAGTTG ACGCCAGTCA 3000  
TCTTTATAAA TAATCTTCATT TCTAGTAATT TCCCAATTT GCTCGTCCAC ATCTATGTCC 3060  
5 GATATATTGT TGAACAAATC CATTAAATCG TTCAATGCCT CAACAGTAAA ACTTTCCCTT 3120  
TTAACTGTAT AAGTTAAAAA TGTCCCATTA TTATCAGTTG AAATTAANAAC TTCAGGTAAT 3180  
ACAAAATGAT TTAGTCCAAA CTCTCGCCAT TCATCATCTG ATTTATGACT TGAANAATGG 3240  
10 AACCTCCAAA CAACTCGAAG ATGATGTTTC TCAGATTGCG GATGTATAAA TGTGATGTTA 3300  
TGTTTTAATT TTCCCCAGTC TTTAAAAATA GATTGTTTAT TTTTAGAATT ATTTTGAAT 3360  
AATTGAATTG CTTTGTAGCC AAAATATGAC GTTCGATTAT CATTCAAAAC CATATAAAG 3420  
15 CGATCTCCTG CCTCATGTTC AGTGAGATGA AATAATGTGC TCGGGTCTAG TGACTGTGAT 3480  
AATTTCACCT CAACTGAAAC CCATTCCCTT GAGCTGCCAT ATATCTCTTT GACAATATCG 3540  
TCCTCTAATA CGCCCGTAGC CATCCATTTT ACTTCTTTCT TCGTCTTTTT TCACTCATTA 3600  
TTATATTGTA TCATTTTTGG ATAATTGTGT TACAAGAATT GCTTAAACTT ATCTTGCAAT 3660  
TTTTCACGTC AATTGACCTT TATGCTACTT TCTATTAAAA TATCTTTGTT ATAAAAATA 3720  
25 TGAITTAAGG AGGTTTGTIA TTCAATGAGT AATCAATATC AGCAATATTC TACAGTTAAG 3780  
AAATATTGGC ATTTAATGGG TCCTCATACA TTAAC TGCTT COGTAGTACC CGTTTATGTT 3840  
GGTACAGCAG CATCTAAAT ATATTTTCTT GGTAGCGAAG ATCATATTAA AATCAGCCTA 3900  
30 TTCATTGCCA TGTTACTAGC ATGCTTACTT ATTCAAGCAG CAACTAATAT GTTTAATGAA 3960  
TACTATGATT ATAAAAAGG CCTGATGAT CATGAATCTG TAGGCATTGG TGGTGCCATT 4020  
GTTGCAACG GTATGAGCCC AGAGCTTGTG CTACGATTAG CAATTGCATT TTACATCTTA 4080  
35 GCAGCAATAT TAGGTTTGTT TTTAGCTGCT AACTCTTCAT TTIGGTTATT ACCAGTTGGA 4140  
TTAGTATGTA TGGCTGTGG TTACCTATAT ACAGGTGGCC CTTTCCCTAT TTCATGGACG 4200  
CCTTTGGGTG AATTATTCTC AGGCGTATTT ATGGGTATGT TTATTATCGT TATTGCATTC 4260  
40 TTTATTCAAA CTGGCAATAT TCAAAGTTAT GTAATTTGGT TAAGTGATACC TATAGTAATC 4320  
ACTATCGG 4328

45 (2) INFORMATION FOR SEQ ID NO: 280:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1450 base pairs  
(B) TYPE: nucleic acid  
50 (C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

55

|    |  |      |
|----|--|------|
|    | GTTCATACAC GAAAAAATAA ATTTAGATGT TGAAGCATCC TACAATTAAAT ACAGATCCAT | 60   |
|    | TTCAATATAT TTAACATAAA ATCTCGGGAT TTCTAAATTT TGAAATTCG AGGTTTTAT    | 120  |
| 5  | ATTTTTATT AAAATAGCAC ATTTATACTT TATAATAGTA AAGATGAACA TATAAGGAGG   | 180  |
|    | CCAAATCATG GCAAAACATC CATTGGAACA ATTTAATCTA GAATCTAGTT TAATTGACGC  | 240  |
|    | TGTGAAGAC CTTAATTTTG AAAAACCAAC TGAAATTCAG AATCGAATTA TTCCAAGAAT   | 300  |
| 10 | ACTAAAGAGA ACAAATTTAA TTGGTCAATC TCAAAOOGGT ACAGGGAAAT CTCATGCATT  | 360  |
|    | TTTATTACCA TTAATGCGT TAATTGATAG TGAAATAAAA GAACCACAAAG CAATCGTAGT  | 420  |
|    | TGCACCAACA AGAGAAGCTT CACAACAAC ATACGATGCA GCGAACCATT TAAGCCAATT   | 480  |
| 15 | TAAAGCTGTT GTTTCAGTTA AAGTTTTTAT TGGTGGTACA GATATAGAGA AAGATAGACA  | 540  |
|    | ACGTTGTAAT GCACAACCAC AATTGATTAT AGGCACCCCT ACTAGAATTA ATGACTTAGC  | 600  |
| 20 | TAAAACGGGA CATTTACATG TGCATTAGC ATCATATTTA GTTATTGATG AAGCGGATCT   | 660  |
|    | TATGATTGAC TTAGGATTAA TTGAAGATGT AGATTACATT GCTGCAAGAT TGGGAAGATA  | 720  |
|    | TGCAATATTT GCGGTGTTTA GTGCTACAAT CCCACAACAG TTACAACCAT TTTTAAATAA  | 780  |
| 25 | ATATTTAAGT CATCCAGAAT ATGTAGCTGT CGACAGTAAA AAACAAAATA AAAAGAACAT  | 840  |
|    | CGAATCTAT TTAATACCTA CTAAGGTGC AGCTAAAGTT GAAAAGACTT TAAATTTAAT    | 900  |
| 30 | TGATATACTA AATCCATACT TATGTATTAT TTTCTGTAAT AGTAGAGATA ATGCAATAGA  | 960  |
|    | TTTAGCACCT TCACTAAATG AAGCTGGTAT TAAAGTTGGT ATGATTTCAT GTGGCTTAAC  | 1020 |
|    | GCCaCGTgAA CGTAACAAC AAATGArCG TATACGTAAT TtaGAATTCC aATACGTTAT    | 1080 |
| 35 | TGCCaGGGAT TTAGCATCTC GTGGTATTGA TATTGAAGGT GTTAGTCrTG TCATCaATTT  | 1140 |
|    | TGATGTGCCA AATGATATTG ACTTCTTTAC GCATAGAGTC GGACGAACGT GTCGTGGGAA  | 1200 |
|    | TTATrFAGT GTAGCAATTA CGCTTTATAG TCCTGATGAA GAACACAATA TTTCATTAAT   | 1260 |
| 40 | AGAAATATCG GGTTTTGTAT TCAATACTGT TGATATTAAA GATGGTGAGT TAAAAGAAGT  | 1320 |
|    | TAAAGCGCAC AATCAGCGTC AAGCAAGAAT GCGCAAAGAT GACCATTTTAA CTAATCAAGT | 1380 |
|    | GAAGAACAAA GTTCGAAGTA AAATTAAAAA CAAAGTTAAA CCAGGTTATA AGAAGAAATT  | 1440 |
| 45 | TAAACAAGAA   | 1450 |

(2) INFORMATION FOR SEQ ID NO: 281:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1139 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

EP 0 786 519 A2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 281:

|    |  |      |
|----|--|------|
|    | AGTCAGGIAT ATCATGCCaT yCTGAATTGG TCGATATTAA TATCAGTGGT GTTAAAGAAC  | 60   |
| 5  | GAATTGTATA CCAATAGACG CTTTATATTG TAAATAGTA TTAATGCaG AATGAGAGG     | 120  |
|    | AGATTTAATG CGATATGACA AATTATAAAG TTGTCGTTTT AGATATGGAT GACACATTGC  | 180  |
|    | TAAATTCAGA TAATGTGATA TCAGAAGAAA CTGCAAAATTA TTTAACAGCA ATTCAAGATG | 240  |
| 10 | AAGGTTATTA TGTGTCTCTA GCATCTGGTA GACCTACTGA AGGTATGATT CCACTGCTA   | 300  |
|    | GAGATTTAAA ATTACCTGAA CATCATAGCT ATATTATTAG TTATAACGGT AGTAAACGA   | 360  |
| 15 | TTAACATGAC TAATGAAGAA GTAGAAGTAA GTAAATCGAT TGGTAAGCAA GATTTCGATG  | 420  |
|    | AAATTGTAGA TTTATGTGGA GATAGAGGCT TTTTCGTTCT TACATATCAT GATGGTCAAA  | 480  |
|    | TTATTTCAGA CAGCGAACAT GAGTATATGA ATATTGAAGC AGAATTAACA GGTTTACCGA  | 540  |
| 20 | TGAACCGTGT TGATGATATC AAAGCGTATA TTCAGGCCGA TGTACCCAG GTCATGGGTG   | 600  |
|    | TAGATTATGT AGCGAATATT ACAGAAGCTA GAATTGATTT GAATGSGTGT TTCATGATA   | 660  |
|    | ATGTAGATGC TACGACAAGT AAGCCATTCT TCTTAGAATT TATGGCCAAA GACGTTTCAA  | 720  |
| 25 | AAGGTAATGC AATTAAAGCG TTATGTCACA AATTGGGATA TTCGSGTGAT CAAGTCATTG  | 780  |
|    | CTTTGGTGA TAGTAGAAT GATAAATCAA TGTTTGAAGT CGCAGSTCTA GCTATTGCTA    | 840  |
|    | TGGGGaATGC ATCAGATGAA CTTAAGCAAT ATGCAAATGA AGTTACGTTG GATCATAATG  | 900  |
| 30 | AAATGGTAT TCCACATGCG CTCAAAAAAT TGTATATAAT TTTAAATAA GCCTTAACAC    | 960  |
|    | ATGATATTTG AATAAGATAT CTTGTGGTTA AGGCTTTTTA TTTTGTGAA AATGACTTCA   | 1020 |
|    | GTTATACTAT GGAGGATTG AAATACATAT TTTAGATTAG TAATGATATC AAACGAATAG   | 1080 |
| 35 | AGTAAATGTA TATTTtTGa ATAAATCAAG TATTAAGTAG TCACGGAAGG nAGATAAAT    | 1139 |

(2) INFORMATION FOR SEQ ID NO: 282:

|    |                               |
|----|-------------------------------|
| 40 | (i) SEQUENCE CHARACTERISTICS: |
|    | (A) LENGTH: 2931 base pairs   |
|    | (B) TYPE: nucleic acid        |
|    | (C) STRANDEDNESS: double      |
|    | (D) TOPOLOGY: linear          |

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 282:

|    |   |     |
|----|---|-----|
|    | TCTAAAAATG CTGTGAaATT CTTTATATAA TATCTAAAG GAATTAAATG TGATAACATT  | 60  |
| 50 | GCTGTGATAG GAAGTAAGAC AGCGCAATAT TGTGAATCAC TTGGCATTCT AGTTGATTTT | 120 |
|    | AtGCCAAAGC ACTTTCTTCA AGAAGGATTT TTAaATCAT TTAATCAAC TAACCAAAAA   | 180 |

|    |   |      |
|----|---|------|
|    | AATGAAGTTG TTAATAATAGA TTTATATACT TCAGTGCCTA ACAAACAAAA TATACAAGAT  | 300  |
|    | GTTAAGAAAA TGATAGAACA TCAACAAATC GATGCATTAA CATTTTCAAG TTCCTCGGCA   | 360  |
| 5  | GTACGTTATT ATTTTAATGA AGGATTGTGA CCAAAATTC AAGTCGTATT TGCTATTGGA    | 420  |
|    | GAACAAACAG CACGGACCAT TAAATCATAT CAACAACCAG TAACAATTGC AGAAATTCAA   | 480  |
|    | ACACTCGAAT CACTAATGA AAAGATTTTA GAAAGTAGGG GCTAAAAATG AAATTGTGATA   | 540  |
| 10 | GACATAGAAG ATTGAGATCA TCAGCGACAA TGAGAGATAT GGTTAGAGAG AATCATGTAA   | 600  |
|    | GAAAGAAGA TTTAATATAT CCAATTTTTG TAGTTGAAAA AGACGATGTG AAAAAAGAAA    | 660  |
| 15 | TTAAGTCATT GCCAGGTGTA TACCAAAATCA GTTTGAATTT ACTTGAAAGT GAATTAAGAAG | 720  |
|    | AAGCTTATGA CTTAGGCATA CGTGCCTAAT TGTTTTTTCG TGTTCACAAAC TCAAAAGATG  | 780  |
|    | ATATAGGTAC TGTGTCATAC ATTCAAGATG GTGTTATTCA ACAGGCAACA CGTATTGCTA   | 840  |
| 20 | AAAAAATGTA TGATGACTTA TTAATTGTTG CAGACACTTG TTTATGTGAA TATAGTGATC   | 900  |
|    | ATGGTCATTG TGGCGTGATT GATGACCATA CACATGACGT TGACAATGAT AAATCATGTG   | 960  |
|    | CACTACTTGT TAAACAGCA ATTTCTCAAG TGAAGCTGG TGCTGATATT ATTGCGCCAA     | 1020 |
| 25 | GTAATATGAT GGATGGTTTT GTTGCTGAAA TTCGTGCTGG ATTAGATGAA GCCGCTATT    | 1080 |
|    | ACAATATCCC TATAATGAGT TATGGTGCA AGTATGCATC AAGTTTCTTT GGACCTTTTA    | 1140 |
|    | GAGATGCAGC AGATTCAAGC CCATCATTTG GGGATAGAAA AACGTATCAG ATGGACCCCTG  | 1200 |
| 30 | CTAACCGTTT GGAAGCACTT CGTGAATTAG AAAGTGATCT TAAAGAAGGG TGCACATGA    | 1260 |
|    | TGATTGTTAA ACCTGCTCTA AGTTATTAG ATATAGTTTG AGATGTTAAA AATCATACGA    | 1320 |
|    | ATGTTCCAGT TGTGTCATAT AATGTGAGTG GAGAATATAG TATGACTAAA GCAGCGGCAC   | 1380 |
| 35 | AAAATGGTTG GATAGATGAA GAACGTGTCG TTATGGAACA AATGTTTCA ATGAAACGTG    | 1440 |
|    | CAGGTGCTGA TATGATTATT ACGTATTTTG CAAAGGACAT TTGTCGCTAT TTAGATAAAT   | 1500 |
| 40 | AAGGTTTTAT ATTTATGATT TTCCATAAAC TGTAGGAGGA ATTTACTTTA TGAGATATAC   | 1560 |
|    | GAAATCAGAA GAAGCAATGa AGGTTGCTGA AACTTTAATG CCTGGTGGTG TAAATAGTCC   | 1620 |
|    | AGTACGCGCA TTTAAATCAG TAGATACACC AGCAATTTTT ATGGATCACG GTAAAGGTTG   | 1680 |
| 45 | AAAAATTTAT GATATCGATG GTAACGAGTA TATCGACTAT GTACTAAGTT GGGGACCACT   | 1740 |
|    | TATTTTAGGA CATAGAGACC CTCGAAGTTAT TAGTCATTTA CATGAAGCAA TTGATAAAGG  | 1800 |
|    | TACAAGTTTT GGTGCATCAA CATTACTTGA AAATAAATG GCGCagCTCG TTATTGACCG    | 1860 |
| 50 | AGTACCTTCA ATAGAAAAAG TGCGTATGGT GTCATCTGGT ACAGAAGCTA CATTGGATAC   | 1920 |
|    | TTTAAGATTA GCACGTGGTT ATACTGGCAG AAATAAAATT GTGAAATTTG AAGGTTGCTA   | 1980 |
| 55 |   |      |

|    |   |      |
|----|---|------|
|    | GCGGATTCT CCTGGTGTGC CTGAAGGTAT TGCTAAAAAT ACAATTACAG TTCCATACAA  | 2100 |
|    | TGATTTAGAT GCACCTAAAA TCGCTTTGCA AAAATTGGa AACGATATTG CTGGTGTAAAT | 2160 |
| 5  | CGTAGAACCT GTTGCTGGTA ATATGGGTGT CGTACCGCCG ATTGAAGGTT TTTTACAGGS | 2220 |
|    | ATTAAGAGAT ATTACGACTG AATACGGCGC ATTGCTAATT TTCGATGAAG TAATGACTGG | 2280 |
| 10 | TTTCAGAGTC GGTATCATT GTGCACAAGG TTACTTTGGT GTGACACCAG ATTTAACTTG  | 2340 |
|    | CTTAGGAAAA GTTATCGGTG GAGGACTACC TGTAGGTGCA TTTGGTGGTA AAAAAGAAAT | 2400 |
|    | CATGGATCAT ATAGCACCAT TAGGAAATAT TTATCAAGCG GGTACGTTAT CAGGAAATCC | 2460 |
| 15 | TCTTGCAATG ACAAGTGGTT ATGAAACGTT AAGCCAATTA ACGCCAGAGA CATATGAGTA | 2520 |
|    | TTTTAATATG TTAGGCGATA TACTTGAAGA CGGTTTAAAA CGTGTATTTG CTAAACACAA | 2580 |
|    | TGTACCAATA ACTGTAATA GAGCAGGTTC AATGATTGGT TATTTCTTAA ATGAAGGACC  | 2640 |
| 20 | TGTAACATAA TTTGAACAAG CGAATAAAG TGATTTGAAA TTATTTCGAG AAATGTATCG  | 2700 |
|    | AGAAATGGCA AAAGAAGGTG TGTTTTTACC ACCATCTCAA TTTGAAGGTA CATTCTTATC | 2760 |
|    | TACGGCACAC ACGAAAAGAG ATATTGAAAA AACGATTCAA GCATTGATA CGGCTTTAAG  | 2820 |
| 25 | TCGTATTGTA AAATAAATAT ACGGACAAAT TGAGAGCCTG AACTTTGTTT AGGCTCHTTT | 2880 |
|    | TAAATGTATA TAAGGCATGG GCGGCGACTT GATAGTGAAA GTCCACTACT A          | 2931 |
| 30 | (2) INFORMATION FOR SEQ ID NO: 283:                               |      |
|    | (i) SEQUENCE CHARACTERISTICS:                                     |      |
|    | (A) LENGTH: 1421 base pairs                                       |      |
|    | (B) TYPE: nucleic acid  |      |
|    | (C) STRANDEDNESS: double  |      |
| 35 | (D) TOPOLOGY: linear  |      |
|    | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 283:                        |      |
| 40 | AATTATGAAT GCATTACCAG TATTATTACA AAAGAACAAT TAAAAATGTT TGTATTGAT  | 60   |
|    | TATGATACGC ATCTCATTAA AAATGTAaTG GTTGCCAGCAG ACGTGTAAAA GGCAATGAT | 120  |
| 45 | ATTCAGGAC ATGAACCATT AATCGTTAAC CTTCAAACGA TTGATGAaAC ATTACATCGT  | 180  |
|    | TTACCTATGC ATAATAGAAA AGACATGATG GTTAATGGCG GTGACTTAT GGCACATTTA  | 240  |
|    | AATGCCAAAA GTGGTCCGTG GTTAAAGAT GTGCTAAGAC AAATTGAGAT AGCGATTGTA  | 300  |
| 50 | ACAGGTAAGG TAAGCAACGA AGAAACTGAA ATTTTGAAAT GGGTGGATAA TCATGTCAAA | 360  |
|    | ATATAGTCAA GATGTACTTC AATTACTCTA TAAAAATAA CCGAATTATA TATCTGGACA  | 420  |
| 55 | AAGCAATTCG GAATCACTTA ATATTTCACG CACTGCAGTA AAAAAAGTGA TTGACCAATT | 480  |

EP 0 786 519 A2

CCCAGATAIT TGGTATCAAG GTATAATAGA CCAATATACA AAAAGTCTG CTTTGTGTTGA 600  
 5 TTTTAGTGAA GTATACGATT CAATAGATTC TACACAACCTT GCTGCGAAAA AGTCACTTGT 660  
 TGGAAATCAA TCITCATTTT TTATCTTGAG TGATGAACAA ACGAAAGGTC GTGGCGGATT 720  
 TAATAGACAT TGGAGTCTTT CAAAAGGGCA AGGACTTTGG ATGTCTGTG TGTTAAGACC 780  
 10 TAACGTTGCA TTCTCAATGA TATCTAAATT TAATTATTT ATTGCATTAG GGATAAGAGA 840  
 TGGCATTCAA CATTTTAGTC AAGATGAAGT CAAAGTGAAA TGGCCGAATG ATATATTIAT 900  
 TGATAATGGT AAAGTGTGTG GTTCTTAAC TGAAATGGTT GCTAATAATG ATGGTATAGA 960  
 15 AGCAATAATA TGTGGTATAG GTATTAATTT GACGCAACAA CTAGAAAACT TTATGAAAG 1020  
 TATTAGACAT AGAGCAACAA GTATACAATT ACATGATAAA AATAAATTAG ATAGATATCA 1080  
 ATTTTAGAG ATATTACTTC AAGAAATTGA AAAAAGATAT AATCAATTTT TAACGTTACC 1140  
 20 TTTTCTGAA ATTGCTGAAG AATATACTGC AGCTCTAAT ATTGGAATA GAACGTTGCT 1200  
 ATTTACAGAA AATGATAAAC AGTTTAAAG ACAAGCAATT GATTAGATT ACGATGGCTA 1260  
 TCTAATTGTT AGAGATGAAG CGGGTGAATC ACACCGTTTA ATTAGTGCAG ATATAGATTT 1320  
 25 TTAACACTAA AGCAAGSAGA GATAGCTATG GGTATGGCAA CCTATGCCGT TGTGATTG 1380  
 GNAACACAG GCAACCAATT AGATTTTGAC GATATCATTC A 1421

30 (2) INFORMATION FOR SEQ ID NO: 284:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 2202 base pairs  
 (B) TYPE: nucleic acid  
 35 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 284:

40 CCAAGTTGCC TAAATGATT AAGCAAGGT TATACCCTAT GCAACGATT GAACAAGAAT 60  
 CTGGAGCCAT CCGACTGCCA ACGATTCTA GAGTGAAGCG TTCATTACAA TGGGGTAATG 120  
 45 ATGCTTATAC AATGATTTTA GATCGTATGA ATATTGAAC AAATGAATAA TAAATGAACG 180  
 ATAAACAATG GTTATCTATC TGCACATAA AGGTAGATAA TCATTGTTTT TTCACGAAAA 240  
 AATTACAGA GTAAAAGAAC TTAATTTC TATTAAGTCT TTGAACCTCG ACACTTAAAA 300  
 50 ATGCTATAAT CATATGTATG TTAATAAAGG AGTTTCGGAA AATGTATGAC ATTAATAAAT 360  
 GGCGCCATAT TTTTAAATTA GACCCAGCTA AACATATTC AGATGATGAT TTAGATGCGA 420  
 TTGTATGTC TCAAACAGAT GCAATTATGA TTGTTGGAaC TGATGACGTT ACTGAAGATA 480

55





(A) LENGTH: 785 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 285:

```

10 AGTGGTGCAA AGATAGGCAT TGATaATACC GCTAAGCCAG CAAGATGATG GCACGATAAA      60
   ACCTAAACAG AAGAAmATAA ATAGTAATAC GATGATAAAT AATGGTCCAC TCATATGTTG      120
   AACTAAAGAT GATGAAAAGT GTAAGATTGT ATCTGAAATC ATACCTTCAT TCAACACTAA      180
15 ATTAATACCT CGAGCTAAAC CAATAATTAA AGATACACCT ACTAAACTTG ATGCACCCATT      240
   GACAAATGCA TCTACAGTTC CTTTTTCTCC CAATCCAGAT TTACCTGTCC CAGCAATAAA      300
   CATTATTATA ATTGTAAATA TAAAAATGCT TGAAGCCATA ACTGGGAACC ACCAACCTTG      360
20 CGTCATAACT CCCCATACCA TAATTGGAAA TGGTAGTACA AATAATGTAA GGATTATCTT      420
   CTTACGCAAA GTAAAAATGGG CACTATCGTC ATCTTTTAAT ACAGACCATT GCTGTTCAAA      480
   AGCATCTTTG TCTTCATAAG AATATGACGC TTAGGATCG TTTTAAATT TTTTACAGTA      540
25 CCAATATAAA TAACTAATAA CAAAAATCGC ACCGACAATA CAAGCACCTA TTCTCCAATA      600
   CAAGCCATCC GTAAAAGTTG TACCAGCGGC ATTAGAGGCA ATTACAACCG AGAAGCGGTT      660
30 AATAGTTGAA AATGTACTAC CGACAGAGCT GGCAGGAAT ATGGCACCAA CTGGAAACGA      720
   TAGAATCGTA TCCTAACGCT AATAAATATA GGGACTAAAA TGGATAAAA TGCTACAGCT      780
   TCTTC                                           785

```

## (2) INFORMATION FOR SEQ ID NO: 286:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 812 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 286:

```

45 CTAACGnGAT AAGGTTGCAA nTTTATCTGA ACATCTGATG ACTGTAAATT TGTaATGAT      60
   AAAATATTG TCACTAATAG ATATAAATAC TGACTTTCCT GAAACTATG TACAAGTAAT      120
50 TGTTCCTTTT CTATGATAGA CATATCTTTA CTATGTGATA CTAATAATATC TAAATKTCCT      180
   ATAATTGTTG TTAACGTTGT ACGTATGTCA TCGGAAATTG ATCTTAAAAA ATTGTAATGT      240
   GTCAGTTGAC GTTCAGCCTG TAACATGGAT TCTCTGTTT GTTTAAGTAA CGTCACATTT      300

```

|    |            |             |             |            |            |             |     |
|----|------------|-------------|-------------|------------|------------|-------------|-----|
|    | ATCAGTTGAG | AACCTTTGGTA | ATCAATGGCT  | AGAAATGCCT | TAATCGGAGA | TGTGCCCAATT | 420 |
|    | GGTATCAACC | ATTATTAAAT  | GCCTGGGAAT  | GTATCTGGTG | TGCAACACAG | TTGCTTTTCA  | 480 |
| 5  | TTTTTAATTA | CCCAGCTTAA  | TGCTTGTTCA  | TGCTGTTGAG | TGTAATTATC | GATATGGTTT  | 540 |
|    | TGCAATGGTA | TGTTTTAAT   | TACTTTCGAT  | TGATTGTAA  | CGTATATAGT | AATTGATTGT  | 600 |
|    | TGCAATAATT | GATTAAATTG  | GATATCCAGCA | TTTATTAGTA | AGTTTTCAAC | TGTATAAGTT  | 660 |
| 10 | TGTTTAATCG | AATCATTAAG  | TTGAAATTAAT | AAATCTGTAC | GATAAAGTGT | CTTTTTAGTA  | 720 |
|    | ATGGAGTAT  | GGAATTTAAT  | TGTTTAAAT   | AAAGCACTCG | TTAAATAATC | TGTTAAAAATG | 780 |
|    | CTAAGCATAA | AGTTAATAGG  | ATAGTCAAGG  | CG         |            |             | 812 |

(2) INFORMATION FOR SEQ ID NO: 287:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1732 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 287:

|             |             |              |            |            |             |     |
|-------------|-------------|--------------|------------|------------|-------------|-----|
| ATTnATATTAT | ATTACTGCTA  | TTTTAAATTT   | TAAAAAATGC | TTTTGATTAT | ATTCAACAnT  | 60  |
| TTGTATAAAA  | TTAAATTTGC  | TTTGTATTAA   | AGCATGAAAA | TTGTAATCAA | ACCATAAAAT  | 120 |
| GTCGTATGAT  | GTAGTTAGAA  | TTTTAAATG    | CAGGAGGTCa | AGTATATGAC | TGAATAaACA  | 180 |
| TTCAAAGGTG  | GACCAATCCA  | CTTAAAGGT    | CAACAAATTA | ATGAAGGTGA | TTTTGCACCT  | 240 |
| GATTTTACAG  | TGTTAGATAA  | TGACTTAAAT   | CAAGTAACAT | TAGCAGATTA | TGCTGGTAAA  | 300 |
| AAGAAATTAA  | TTAGTGTGTT  | ACCATCAATT   | GATACAGGTG | TTTGTGATCA | GCAGACTCGC  | 360 |
| AAATTCACAT  | CTGATGCTTC  | TAAAGAAGAG   | GGGATTGTGC | TTACAATTTT | AGCAGACTTA  | 420 |
| CCATTTCGCAC | AAAAAAGATG  | GTGCGCTTCA   | GCAGGTTTAG | ACAATGTCAT | TACATTAAGT  | 480 |
| GACCACCGTG  | ACTTATCATT  | TGGTGA AAAAC | TATGGCGTTG | TTATGGAAGA | ACTTCGCTTA  | 540 |
| TTAGCTCGTG  | CAGTATTGTT  | ATTAGATGCA   | GATAATAAAG | TGTTTATATA | AGAAATCGTT  | 600 |
| AGTGAAGGTA  | CTGATTTCCT  | AGATTTTGAT   | GCTGCTTTAG | CTGCATACAA | AAATATTTAA  | 660 |
| TCATTATAAGA | GATAAATCTT  | AAAATGTATA   | CATGTGTGCC | ATCGTTGTCA | ACGACATTA   | 720 |
| AATAGAAATG  | TTTCTTATTA  | TTCCTAAGAC   | CTATGGGCAc | TTTTTATTGG | AGAGGGACGA  | 780 |
| ATTATGGCAGA | ACACCAAAACA | ATTATGGAAc   | GCTTGTTTCA | TACATTAGAT | GA AAAAGCTA | 840 |
| AAACATTAAA  | TAATGAAAAAT | GGCCaAAAGTT  | TTATTGAAAA | TCTTGGGCTA | GCAATTGGAAC | 900 |

|    |   |      |
|----|---|------|
|    | CATTCCAATT TGCAATTTTA AGTTTAATGC aGGAAGAAAA GATAcAAGCA AATCATCAAA                                       | 1020 |
|    | TTACACCAGA TTCaATTGGA TTGATACTAG GATTTTTAGT TGAGCGTTTT ATGAACAACC                                       | 1080 |
| 5  | AAGAAGAATT ACATATTGTT GATATTGCAA GTGGTGCCGG TCATTAAAGT GCTACTGTAA                                       | 1140 |
|    | AAGAAGTGTT ACCTG <del>ra</del> AtT GcGGTTATGc ATcATT <del>Ta</del> T TGAAGT <del>TG</del> AT CCAGTTTTAT | 1200 |
|    | CACGTGTTAG TGTACATTTA GCAAACTTCT TAGAAATCC TTTGATGTG TATCCTCAAG   | 1260 |
| 10 | ATGCCATCAT GCCACTACCA TTAGAAGAAG CAGATATCGT TATTGGTGAT TTTCCAGTAG                                       | 1320 |
|    | GCTATTATCC AATTGATGAA AGAAGTAAGG AGTTTAAGCT AGGTTT <del>TG</del> AA GAAGGACATA                          | 1380 |
| 15 | GTTATTCA CA TTTTATTATA ATAGAACAAG CAATAAATGC ATTA <del>AA</del> AAGAT GCTGGATATG                        | 1440 |
|    | CCTTCTAGT GGTACCAAGT AATATTTT <del>TA</del> CAGGTGAACA TGTA <del>AA</del> ACAG CTGAAAAAT                | 1500 |
|    | ATATTGCAAC AGAGACAGAG ATGCAAGCAT TTTTAAATTT ACCACCAACT T <del>TA</del> TTTAAAA                          | 1560 |
| 20 | ATGAAAAAGC GCGAAATCT ATATTAA <del>TT</del> T TACAAAAGAA AAAATCGGT GAAACAAAGC                            | 1620 |
|    | CAGTTGAAGT ATTATTGGCA AATATTCCTg ATTTcAAAA TTCCTTCACC AATT <del>TC</del> CAAG                           | 1680 |
|    | GATTTATGGA CAGAGTTAAA ATCCAGTGGG ATGGGACCAC CAAATCGTCC TA   | 1732 |
| 25 | (2) INFORMATION FOR SEQ ID NO: 288:   |      |
|    | (i) SEQUENCE CHARACTERISTICS:   |      |
|    | (A) LENGTH: 2779 base pairs   |      |
|    | (B) TYPE: nucleic acid  |      |
| 30 | (C) STRANDEDNESS: double  |      |
|    | (D) TOPOLOGY: linear  |      |
|    | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 288:  |      |
| 35 | AAAAGAACTA GCTAAACGCA AGCAAGAAGC TATTAGTAGA ATTAAGAACT TTTCAAAATGA                                      | 60   |
|    | AAAA <del>A</del> TAAAT AGTATTGGA ATAGTGAAAT TGGCACAGCT GATGAAAAAC AAGCAGCAAT                           | 120  |
| 40 | GAATCAAATT AACGAAATTG TGCTTGAAC AATTAGAGAT ATTAATAATG CGCATACTT   | 180  |
|    | ACAGCAAGTT GAGGCTGCAT TGAACAATGG TATTGCTCGA ATTTCAGCAG TACA <del>AA</del> TTGT                          | 240  |
|    | AACATCTGAT CGTGCTAAAC AATCGTCAAG TACTGGAAAT GAATCTAATA GCCATT <del>TA</del> AC                          | 300  |
| 45 | AATTGGTTAT GGAAGT <del>GC</del> AA ATCATCCATT TAACAGTTCG ACTATTGGAC ATAA <del>AA</del> AGAA             | 360  |
|    | ACTTGATGAA GATGATGACA TTGATCCACT TCATATGGGT CACTTTAGTA ATAATTTCGG                                       | 420  |
| 50 | TAATGT <del>TA</del> TAT AAAAAGCTA TTGGTGTGGT GGGTATCTCT GGTTTACTAG CTA <del>GT</del> TTCTG             | 480  |
|    | GTTCTTCATT GCCAAACGTC GTCGTAAGA AGATGAAGAG GAAGAATTAG AAATAAGAGA  | 540  |
| 55 | TAATAATAAA GATTCAATAA AAGAGACTTT AGACGATACA AAACATT <del>TAC</del> CACTTTTATT                           | 600  |

|    |            |            |             |            |            |             |      |
|----|------------|------------|-------------|------------|------------|-------------|------|
|    | AAATAATGGC | GAGTCACTCG | ATAAAGTTAA  | ACATACGCCG | TTCTTCITAC | CAAAACGTCG  | 720  |
|    | TCGTAAGAA  | GATGAAGAAG | ATGTGGAAAGT | TACAATGAA  | AACACAGATG | AAAAAGTGTT  | 780  |
| 5  | GAAAGATAAC | GAACATTAC  | CACCTCTATT  | CGCAAAACGA | CGCAAAGATA | AAGAGGAAGA  | 840  |
|    | TGTTGAACA  | ACAACTAGTA | TTGAATCTAA  | AGATGAGGAC | GTTCCCTTAT | TATTGGCTAA  | 900  |
|    | AAAGAAAAAT | CAAAAAGATA | ACCAATCCAA  | AGACAAAAAG | TCAGCATCAA | AAAAACTCTC  | 960  |
| 10 | TAAAAAGGTA | CGACGTAAAA | AGAAGAAAAA  | GAAAGCTAAG | AAAAATAAAA | AATAATTTGT  | 1020 |
|    | TTCTTTGATA | AATAGaGGAG | CACCGATTGA  | CATCACATCA | GTCGGTGCTC | CTTTTATTTA  | 1080 |
|    | TTCTTTTAA  | TTAATTATA  | CAATGCCTGT  | TGAGCGTGTT | GATTCGCTTC | TTTGTTTTGT  | 1140 |
| 15 | TCTCTCGTA  | TCCATTAAAC | AAATAATAAA  | TCAAAATCTT | TTTCAAAAT  | TTCTATTGTA  | 1200 |
|    | TCAAAATAAG | GTTTTGAAAT | TGCGTITTTT  | ACAATACCAG | CTTCAATGCT | ATCTCGCAAT  | 1260 |
| 20 | AGCTTTTAGT | CTGTATAATA | TAGTGGGTAT  | TGAACATTTA | ATTACAGTGC | ATGTCTTAAGT | 1320 |
|    | GCATAAATAC | ATGCAGCCCA | TTCTGCAGTG  | TGGTTATCCA | TTTCGCCATA | CTCATGTGTA  | 1380 |
|    | TATGTATAAT | GCTGCTCATC | TTCTTTGATT  | ACAATGGCAC | ATGTACTTAT | GCCTGGATT   | 1440 |
| 25 | CTTTCGTGC  | CAGCATCAAA | ATTATTTTTC  | GCCATAATA  | ACCTACTTTC | TATTCAATAC  | 1500 |
|    | TTAGTTAAAG | TTACTATTAC | TGTAATACAA  | AATATGTTGG | GTAATCCATT | AAAAAACAGG  | 1560 |
|    | CATCACTTAA | ATAAGTAACA | CGTGTTTAAA  | ATACTCGCTG | AITCAAAGAT | GATTTCTCAA  | 1620 |
| 30 | TACGTaTACT | GTAATATACT | TCCTAAAAAA  | ATCATCTTCA | GGCTGGGACA | TAAATCAATG  | 1680 |
|    | TTCTATGCTC | TACGATGTTA | TATTGGCAGT  | AGTTGACTGA | ACGAAATATG | GCTGTAAACA  | 1740 |
|    | AGCTTTTTC  | AATTCTAGTC | AGGGGCCCCA  | ACACAGAGAA | TTTCGAAAAG | AAATTCTATA  | 1800 |
| 35 | GGCAATGCGA | GTTGGGGTGT | GGGTCCCAAC  | ACAGAAGATG | ACGAAAAGTC | AGCTTACAAT  | 1860 |
|    | AATGTCaAG  | TTTGGGATGG | GCCCCAACAA  | AGAGAAATTG | GATTCCCAAT | TTCTACAGAC  | 1920 |
| 40 | AATGCAAGTT | GGGGTGGGAC | GACGAATAAA  | ATTTTGGGAA | AATATTATT  | CTGTCCCACT  | 1980 |
|    | CCCTTAAAA  | TTATTCTTTT | GTTGAGTAAG  | TGCGTTAATA | GCCTTGATCT | AACTTATCAA  | 2040 |
|    | TCTTACCTTT | ACGATAAAAT | GATTTAGCAA  | TATATCCAAA | TGGTACATG  | AAAACTTTG   | 2100 |
| 45 | AAGCTAATTT | TAATACGTAC | GTTGTAATAA  | ATATTTCAAA | TACAAMTGTa | CCAGGTAAAC  | 2160 |
|    | TTCCGATAAT | TGCGATTAAG | ACAAATAAAG  | CTGTATCAAT | TATTGCGCTT | AAAAATGTAC  | 2220 |
|    | TACCATAATG | CACGATGAA  | AAACGTTTTA  | TcmGaACTAA | ATACTTTTTT | AATTAGTGAA  | 2280 |
| 50 | AAGATAAATA | CATCAATATG | TTGACCAATA  | ATATATGCGA | CGATTGAGCC | TAAAGCAATG  | 2340 |
|    | CGTGGACAA  | CATCAAGAT  | TGCGTGAAT   | GCTTTTTTGT | CCATATCTTC | TGTCACAGGA  | 2400 |

EP 0 786 519 A2

CAAACGCTC TTTTGCAAC TCTACGCCA TAAATATCGT TTAATATATC TGTGCTAAA 2520  
TAAATAGAAG CAACATGAC ATTACCTAAA GTTGCTGAAA TACCAAAGAT TTCTACAGTT 2580  
5 TTAATCAGTT GTATGTTGGC AATGATTGTG CCAATTGCAA CCCATGCAAT TAAACCTTGT 2640  
TTACCAAAAA AGCGATACAT AAGTACCATA AGCAGCAACG TTGCAATAAA CGTAACTAGT 2700  
CCTAAAAATT CATTTATACAT ATTAAAAATG CCTCCTAAAT TTGTATCATG CGGGTGTTTA 2760  
10 GAAACGCTC AATAAATAA 2779

(2) INFORMATION FOR SEQ ID NO: 289:

(i) SEQUENCE CHARACTERISTICS:  
15 (A) LENGTH: 1999 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 289:

ACTGATGTGC GTTCATCAAA AACAATATAA TCAAAATCAT TTTCATCAAA TTGCTTAAAA 60  
25 TTATCATCTC TAGATAATGT TTGAATGGTT GCAAAATAAT ATTTGGCATC GACATCTCTA 120  
TGTTTTCTCG TCAACAATCC AAAATCACTA TCATTTTTTA TAGGTAATAC TTTTTTAAAT 180  
TCTTCTTAG CTCTATTTAA AATCCCCTCA TTATGAACAA TAAATAAAAA TTTATTAGGG 240  
30 TTTACTTCTC TAACATCTAA TGCACATAAA ATCGTTTTAC CTGTACCAGT TGCAGATATT 300  
ATTAACGCCT TATCTTTGGC TTTATCCCTA ATAGCTTTTA ATGACCTTAA TGCTTCTGCT 360  
TGCAATTAAT TGGGTACAAT TTCACTGAT TTTTTCACCT TATCAGCTAG CAGCAITTGA 420  
35 GTTTGTCAA CCTCGCTAA TTTTCTAAG GAACGGTACT CAAATGATTC TTTATATGAA 480  
TTAATCCATT GCTCAGTCAG TGGGGTACTC TTTTGCCATA ACAAGTCAAA TTCACTTTTT 540  
40 ACACATCAAA CTAATCGCC ATTTTTCATA GTAGACAGTA AAACATTATG CTCATAATTA 600  
ACCTTTAAGC CATTAGATGT TAAATTAGAG CTTCTATTA CCATAGAAGT ATAATCCTTA 660  
TGCTCAAAAA TATATCCTTT GGCATGGAAT CCAGCAATAT CAGTTAATCT TACCTCTACA 720  
45 TTTTTTAAAT TAAGTAATTC TCCATACATT TTAGGACTAT TAAACCTTAA GTAATTAGAT 780  
GTTAATATTT TCCCTTTAAC ACCCTTATTG CTTAAATCTA ATAGTTGAGC CTTTAAAGCTG 840  
GCTAAACCGC TTTCTGTTAT AAAAGCCACA GAAAAATAAA ACGTTTCACA TTTTGAAGT 900  
50 TCATCTATAA TTGTTGAAAG AACTTTTTCA TTTTATTAT TACTAAAAG CTCGGTGTA 960  
TAATTCCCTT TATGAGAAAT ATGTTTGTCT ATAAACCTT TATGTAAGA TTGATTGAAA 1020

55

## EP 0 786 519 A2

CGCTGGGGCC CAATTTAATT TATCAAGTTC GTTTATTGAC AACCAITCAA TACTCTTATG 1140  
 TTCAGTITAGA GTTGGTAACT CTTTGTITAA AGTACATTG TATGTGTGTA ACCTAACAAAT 1200  
 5 TCCAAAATCA TATTCATGTT CTGTAGTTAT AACTTTGTCT CCAACAATTA AATCACATTT 1260  
 CATTCTCTCT CTAATTTCTC TAATCAAAGC GTCTTTTCCA GTTTCATTCT TTCAACCTT 1320  
 AChGCCAGGA AATTCCACCA TTAAGGCAG ACTCATTITT TCATTCTCT GTGCACAAAG 1380  
 10 AATTTTGTGA TCAGAAAAAA TAATAGCTCC TACTACATTG ATTACTTTTT TCATAAGACT 1440  
 CACCCTTCAA TTAAAAATCA TCTTAATTGT TATTCTATCA AAAATTACAA AACTATATAT 1500  
 AAATCAATAT TAAAAATTAA TATTTTACAT TCACATGAAC GCTCTACTCC ATGCATTTTC 1560  
 15 ATACACATCT ATTATATAAT ACTTGTGAAA AGTATTGTCT TGGGCTGTG TTTTITTTACT 1620  
 TTTGGGCGCT ATTTCTTTAT AATTCATTAC ATAAATGTAA GGGCTTAGT TTTCATGTTT 1680  
 20 TATTAAGTCT AACTGAGATT TTGAAAGGAT GTTTAGCAAC AATGGATAAA GAATTATGGA 1740  
 TAGAACGAGC TAATGATAGT TTAGTTAAAC ATTTTATGA GCAGCAATCT GACATTGAAC 1800  
 AGCGAGAAAGG TTTTGAAAGT AAATTAACT TGTGACTGC GGGTATACGC GGAATTCG 1860  
 25 GTCTTGTTGA AGGTCGACTT AATAAGTTTA CTATTGAAAA ATTGGCATTG GGTTTAGCGC 1920  
 GTTATTTAAA TGCCCAAACA AACAGTCCAA CAATAGTCAT TCATTATGAT ATTAGACATC 1980  
 TTTCCAACCTG AATTCGCCC 1999

## (2) INFORMATION FOR SEQ ID NO: 290:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1933 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 290:

GATGACTTTT CCCCTCATAT ATCTTCATGG TCCAGGCGTC CATTAAATGCG TCAAAGGATG 60  
 GCACATTTTA CTGGAACAA ATGATTTCATA TGGTTCATAA AAATCACGCG TCGTAATATA 120  
 45 ATCTTCTAAA TCAAATGCAT AGAAAATCAT TGGCTTTTAA AATACTGCAT ATTCATATAT 180  
 TAAAGATGAA TAGTCACTAA TTAATAAATC TGTATGAAC AGTATATCAT TAACTTCTCT 240  
 50 AAAGTCAGAA ACGTCAACAA AATATTGTTT ATGTTGTCT GCAATATTAA GTCTATTTT 300  
 CACAAATGGA TGCAATTTAA ATAATAACAAC CGCGTTAATT TTTTGGCAAT ATCTTGCTAA 360  
 ACGTCAAAA TCAATTTTGA AAAATGGGTA ATGTGCTGTA CCATGACCAC TACCTCTAAA 420

TTGTTTGATC TGTGTCGCAT AAGCTTCATC AAATAGTACA TCAGTACGTG GAACACCTGT 540  
 AGGCACATCA TTTTCTCTTT TAATACCAAA TGCTTCAGCG TAGAATGGAA TATCGTGTTC 600  
 5 AGATGATACA TAAGCTTTTG TATAGCTACG ATGATTTAAT GAATCAATAA ATGGTCCACC 660  
 CTTTTTACCA GTACGACTAA AGCCAACGTG TTTAAAGCA CCAACGCGAT GCCACTCTTG 720  
 AATAACTTCT TGAGAAGCTC TAAAACGCAC TGTATAATC AATGGGTGAA AGTCATCAAC 780  
 10 AAAGATGTAG TCTGCCTTCC CAAGTAAATA TGGCAATCTA AACTTGTGGA TGATGCCACG 840  
 TCTATCTGTA ATATTGCGCTT TAAAACAGT GTGAATATCA TACTTTTTAT CTAATTTTGT 900  
 ACGTAACATT TCGTTATAGA TGTATTCAAA GTTCCAGAC ATCGTTGGTC TAGAGTCTGA 960  
 15 TGTGAACAAC ACGTATTCC CTTTTTCAA GTGAAAAAT TTCGTGTAT TAAATATCGC 1020  
 TTTAAAAATA AATTGCTTGT TATTAAATGA TTGTTGCGG AAATCTTAC GTAATCTTTT 1080  
 20 ATATTACGA ACGATATAAA TACTTTTAA TCCGAGATC GTTACAACAA CATCAAGGAC 1140  
 AAATTCATTA ACATCGCTAG AAATTCAGG TGTACAGTA TAAACGTTT TCTTCGAAAT 1200  
 GCCGCCTTTT CTAATTCCTT TTAGGTAAGT CTGCAATAAG AAATGATT TACCATTTTG 1260  
 25 TGTTCCTAAT TCGTTGTAAT CTTCTCTTG TTCTGGCTTT AGATTTTGAT ATGCATCATT 1320  
 AATCATCTCT GGGTTTAACT GTGCAATATA ATCAAGTTCT TGCTCATTC AATAAGTA 1380  
 CTTATCTTCA GGTAACTAAT AACCATTATC TAAGATAGCT ACATTGAAAC GACAAACGAA 1440  
 30 TTGATCCCA TCTATTTGA CATCATTCGC CTTTATTGTA CGTGCTCAG TTAATTTTCT 1500  
 TAATACAAAA TTACTATCTT CTAATCTAG GTTTTCACTA TGTCTTCAA CGAATAACTG 1560  
 AACACGTTCC CAATAGATT TATCTATATA TATCTTACTT TTAACCAACG TTAATTCATC 1620  
 35 CTTTCTATT TACATAATCC ATTTTAATAC TGTTTACCC CAAGATGTAG ACAGGTCTGC 1680  
 TTCAAAGCT TCTGTAAGAT CATTAAATGT TGCAATTCA AATTCCTGAC CTTTAAACAA 1740  
 40 CGCTAATTTA nTACAATAT CTGGGTATTG AATGTATAAG TCTACAACAT CTTGGAATC 1800  
 TTTTGAACCA CTGCGACTAC TACCAATCAA CGTTAGTCT TTTTCCAATA CTAGCGTGT 1860  
 ATTAACCTCT ACTGGGAAC CACTTACACC TAACAGTnCA ATGCTTCTCT TGTGGAAT 1920  
 45 GTAATCGATC ATT 1933

(2) INFORMATION FOR SEQ ID NO: 291:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 2049 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 291:

|    |   |      |
|----|---|------|
|    | nGInCGnCA GATATATTGG TGGTCTTTAG TAAGTGTATC AAATTCATCA GATGTCAAAG  | 60   |
| 5  | GCATGTTATC ACCTCCTTAG GTTGATAACA aCATTATACA CGaAAGGAGC ATAAaCAAaT | 120  |
|    | GAACACaAGa TCAGAAGGAT TGCATATAGG CGTCCACAA GTTCTTAGCA AAGCTGATGC  | 180  |
| 10 | TTCTTCATCC TATTTAACGG AAAAGGAACG TAaCTTAGGA GCGGAATAT TAGAACTTAT  | 240  |
|    | TAAAAAAAGT GATTACAGCT ACTTAGAAAT AAACAAAGTT TTCTATGCAT TAGATAGAGA | 300  |
|    | ACTTCAATAC AGGGCGAATA ATAACAAACT TTAaCATTTA TCTAAAGGAG TGATAGAGAT | 360  |
| 15 | GCCAAAAATC ATAATACCAC CAACACCAGA AAACACATAT CGAGGGCAAG AAAAAATTGT | 420  |
|    | GAAAAAGTTA TACGCAACAC CTACACAAAT CCATCAATTG TTTGGAGTAT GTAGAAGTAC | 480  |
|    | AGTATACAAC TGGTGTAAAT ATTACCGTGA AGATAAATTA GGTGTAGAAA ATTTATACAT | 540  |
| 20 | TGATTATTCA GCAACGGGAA CATTGATTAA TATTCTTAA TTAGAAGAGT ATTTGATCAG  | 600  |
|    | AAAGCATAAA AAATGGTATT AGGAGGATTA TCAATGAGC GACACATATA AAAGCTACCT  | 660  |
|    | ATTAGCAGTG TTGTGCTTCA CGGTCTTAGC GATTGTACTC ATGCCATTGC TGTACTTCAC | 720  |
| 25 | TACAGCATGG TCAATTGCAG GATTCCGAAG TATAGTGACA TTCATATTTT ATAAGGAATA | 780  |
|    | CTTTTATGAA GAATAAAAAA ACTGCTACTT GTTGAGCAA GTAACAGTGC AAGATGAGCA  | 840  |
|    | ATTGTCTTAA ATAATTATAT AAGGAGTTAT TAATATGACC TTACAACAAA AAATACTATC | 900  |
| 30 | ACATTTTGCA ACATATGACA ATTTCAATTC TGATGATGTT GTTGAAACTT TTGGGATATC | 960  |
|    | TAAACACAT GC AAAATCCA CACTTTCAA ACTTAAGAAA AAAGGAAAGA TTGCAATGGA  | 1020 |
| 35 | AAGTTGGGGT GTCTGGCGTG TTATTGAATC GCAATTGCAT TTAAGTGTAG TCGAACGTAA | 1080 |
|    | AAAAGAAAT TTAGAAGAAC AATTGGAATT GTTAGCAAGA TTAATGAAC AAAGTGATGA   | 1140 |
|    | CCCTAGAGAA ATAGAAGAAC GTATCAAGTT AATGATTCGT CTAGCTAACC AATTTTAAAG | 1200 |
| 40 | AGGAGTTAAT CAATGGCAGT ATTAGAAGGT ATTTTGAAG AATTAAAACT ATTAATAAAG  | 1260 |
|    | AACTTACGTG TGTAAACAC TGAACATCA ACTGTAGATT CATCAATTGT ACAAGAGAAA   | 1320 |
|    | GTTAAAGAAG CACCAATGCC AAAAGAAGAA ACAGCTCAAC TGAATCAAT TGAAGAAGTT  | 1380 |
| 45 | AAGGAACTT CTGCTGATTT GACTAAAGAT TATGTTTAT CAGTAGGAAA AGAGTCCCTT   | 1440 |
|    | AAAAAGCAG ACACCTCTGA TAAGAAAGAA TTAGAAATA AACTTACGA ACTTGGTGGC    | 1500 |
| 50 | GATAAGCTAT CTACTATCAA AGAAGAGCAT TATGAAAAA TTGTGATTT TATGAATGCG   | 1560 |
|    | AGAATAAATG CATGAAGCTA GATCACTCAA ATAGAGCTCA TGCAAAGCTT AGTGCAAGTG | 1620 |
|    | GAGCAAAACA ATGGCTAAAC TGTCCACCGA GTATTAAGGC AAGTGAAGGT ATTGCAGATA | 1680 |

55



GTCTTAAATA TGAAGGCTTA ACACAGTTTG AGTTTAATAA AGCTTTTCAA AATTATAAGC 1800  
 GAAATCAATA TTACAGTGAA GAGTTGCGCG AATATGTTGa AGAGTACGTA GCTAATGTAG 1860  
 5 AAGAAAAGTA TAACGAGGCT TTGaGTAGAG ATGACGATGT AATAGCTTTA TTTGAAACAA 1920  
 AATTGGATyT AGGTAAATAC GTCCCTGAAT CTTTGGTay TGGTGATGTC AtTATATTTT 1980  
 CAGGTGGTGT ACTTGAAATT ATTGACCTTA AATACGGTAA AGGCATTGAA GTTTCAGCTA 2040  
 10 TAGATAATC 2049

(2) INFORMATION FOR SEQ ID NO: 292:

15 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 942 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 292:

ATGATGTTTC TATATTCGTA TTAGGAAAAC CTGTGGTAT TACAACAAAC GCCCTAAAAT 60  
 25 TACAATACCG CTGCTCTAT ACCAATTGCA ACAACAGTTT TAAC TGAAAT ATCTTGTTTT 120  
 TTCACTTCA TTACTCTTt ACATAAAAAA TTCATTATAT TGATGGTGCT TTAGATAAAT 180  
 30 GAATGTCCTA ATCATTTCCT GTACCAATAT GATATAAATC TGAAATGAG TCTTGATTGA 240  
 CTGCTACACC AATATTACT AGCGAGTTAA CATACACAAG AGGTTACCC ACATTAACAT 300  
 CTGCAACGTA TCGCGCAAT TTAATAATAT TTTGATAGAC TTTCTTATCT TGATGATAAA 360  
 35 TTGTTACCA CAAATTATTA CCATGAACAA TTTCCAAGGA TTTTAAGAAAT GCCAATGGAA 420  
 TATTTTGCCA TAATGACCCA AATCTGATAT CTAAAAATC AATGCTTCCT GTAACAGAAAT 480  
 CCTGATTTTT TGTCACTTCT CTTATTCTTA ATGCCTCAAT ACTATCAACA TTAATTGCCT 540  
 40 GACCGAGACG TTCAAACGCT ATCTTATTTG CAGCTAATCT CGCACCATG TATGCATAAA 600  
 CATCTCTACC ATGAAAAGTA TGACTTTCTT CGAATGAGG CAATCGGCTT TTCATCTCAT 660  
 CAATTTGAT AACTTTTTTA ATACCTTCGT AATGTTTAA ATGACTTAAA GAACATTAT 720  
 45 CAGGTGTAAT AATGTAATGA CTGGAATATG TTAAGCAAGC AATGTCGCC TATCACTACC 780  
 TACACCCGGG TCTACCACTG ATACAAATaC TGTGCTTTA GGCCAGTATT TTACAGTTTG 840  
 ATATAACGA TATGACGCTA CCCAAATGTC ATACGGTGGT ATATCATGCG TTAAGTTTTC 900  
 50 AACACGTATA TCATCATTAC AGTATATGCA ACTCCATACA TT 942

(2) INFORMATION FOR SEQ ID NO: 293:

55



## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 294:

5 TACCACCAAA TAATATATTA GCTGGCATT TAATAACATT TAAATTGTG ATGATATCAT 60  
 CAATAAAATG TTGAAACTTC GTAATTCAC CTCATAATC ATCAATTGCT GCTAATTGCG 120  
 CACTCGATGC TTGCTGATCT AAATTAAAA TATTCGACAT GCGTTGACTA TAATAAACTA 180  
 10 AATGTTCTAT TAAGCCATCG TCACCTTTT CTTTGGTGC TGACATGACA GCGATACGTT 240  
 TCAAAGGATA GTGTGGCGCC AATTTAATG TCATTAAATCC ACCTAAGAC ACACCGGTTG 300  
 CACTGATAGA TTCATAACCT TCATTGACTA AAAATTGGTA AGCTTTCTCA ACTTCTTCCC 360  
 15 ACCAATCATC TACATTATAT GTCATGAAAT CTTTCAACAA TAAACCATGA CCTGGATAAT 420  
 TCGTGCGATA ACACTAAAT CCTTGGTCAT TTAAGTCAGC TGCAAGATGC TTCACATCCC 480  
 GATTGTACC TGTAATGAA TGTAATAATA ATATCGCATG TCATTGTG CTTTTAAAT 540  
 20 ACGTGGACTC GGTGTTTTAA TTCTCATTTT TctaTATAcC TCCACTaGT CTAAGaGKT 600  
 TKGCTAAACG CGTTGCGTC GATGATTAA 629

## (2) INFORMATION FOR SEQ ID NO: 295:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 2817 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 295:

35 TATGAAAGTA ATGAATGGTA ATATTATTAA ATTTGATGGA AAAGTAGATA TTGATAATGC 60  
 AGATTAATATC GGTTTTTTAA TTGAGCATCC TAAATTATAT GATAATAAAT CAGGATTGTA 120  
 40 TAACTTGAAA TTATTGACAC AAGTATTAGG TAAGGGTITT GATAAGCAT ACACAGACAA 180  
 AATTATAGAT GCATTGGTA TGAGACCTTA TATTAAAAAG AAAGTTAAGA AATATTCAAT 240  
 GGGGATGAAG CAAAAGTTAG CAATTGCAGT ATCTTTAaTG AATAAACCTA AATTTTAAAT 300  
 45 CTGGATGAG CCTACAAATG GkATGGATCC AGATGGCTCa ATTGATGTGC TGACTACAAT 360  
 TAAGTCTTTA GKAAATGaC TTGATATGAG AATTCTAATA TCAAGTCATA AGTTAGAAGA 420  
 TATTGAATTA ATTTGTGATA GAGCTGTATT TTAAAGAGAC GGCATTTTTG TTCAAGATGT 480  
 50 AAACATGGAG GAAGGTGTTG CATCTGACAC AACGATAGTT ACTGTTGATC ATAAAGACTT 540  
 TGATAGAaCT GAAAAATATC TTGCAGAGCA TTTCCAATTA CAAATGTG CAAAGCAGA 600

55



TTGATCGTTT TAAATTTTAT AACAGTAAAG CTCACCOCTGA TCTTACCGTT AAAGTGAGAG 2520  
 AAAAGGATAA CATCGTTAAG GGGATAATAT TAGTAAGAGA TGAAAAGATA CATACTAATT 2580  
 5 TTGATGGGGG AATTGGTTCG CCGATAAATA ACGCgATTGA AAATCTTGGa TTCgGATATA 2640  
 AAAGrACaAA AGTTGGcAAT GrtTKctCAT CgGTAAAGTA TATTGATAGA GATAACCAAT 2700  
 TAAAATTAAA CTTACTTTAT CAAGATTTAG AAATTAAACG TATTGAATTT TTTAGTAAAT 2760  
 10 AGCTTTAGGT CTTAAAGTtw TAAAAACGA ATGaaTAATT TTATTGGGAT GAGTGAC 2817

(2) INFORMATION FOR SEQ ID NO: 296:

15 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1607 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 296:

TCTGTTAAAA TGATTTTCT TTTaAAAGG CCGaAAATCA ATGTTCGATT nTTATTGCA 60  
 25 TTATGGTCTC GATATTGGTA GAATATCAAA TGGTTAAATG AGAAAAACTT GGAGGTGCTC 120  
 ACATGTCATC AATCGTAGTA GTTGGGACAC AATGGGAGAG CGAAGGAAA GGA AAAATAA 180  
 CGGATTTCTT GGcAGAACAG TCAGATGTTA TCGCGGTTT TTCAGGTGGT AATAATGCAG 240  
 30 GCCATACCAT TCAATTTGGC GGAGAAACAT ATAAATTACA TTTAGTACCA TCTGGTATCT 300  
 TTTACAAAGA CAAATTAGCG GTAATCGGTA ACGGAGTCGT TGTGTATCCA GTTGCACTAT 360  
 TGAAGAATT AGACGGATTa AATGAACGTG GCATTCCTAC AAGTAATTIA CGTATATCTA 420  
 35 ATCGTGGCGA AGTGATTTTA CCATATCACT TAGCACAAGA TGAATATGAA GAACGTTTAC 480  
 gTGgGACAA TAAGATTGGT ACAACTAAAA AAGGTATCGG TCCAGCATAT GTAGACAAAG 540  
 40 TTCAACGTAT CGGTATTCGT ATGGCAGATT TACTTGAAAA AGAAACATTC GAAAGATTAT 600  
 TAAAATCAAA CATTGAATAT AAACAAGCAT ATTTCAAAGG TATGTTTAA C GAAACATGTC 660  
 CATCATTTGA TGATATCTTT GAAGAATATT ATGCAGCAGG TCAACGCTCA AAAGAATTTG 720  
 45 TAACAGACAC ATCAAAAATC TTAGACGATG CATTGTAGC AGATGAAAAG GTACTTTTCG 780  
 AAGGTGGCGA AGGTGTAATG TTAGATATCG ACCATGGTAC ATATCCATT C GTTACATCAA 840  
 GTAATCCAAT TGCAGGTAAC GTTACTGTTG GTACAGGTGT AGGTCTCTA TTCGTTCAA 900  
 50 AGGTAATTGG TGTATGTAAA GCTTATACAT CACGTGTTGG TGATGGTCCA TTCCCTACTG 960  
 AATTATTCGA TGAAGATGGA CATCATATTA GAGAGGTGG TCGTGAATAC GGTACAACAA 1020

55

TAAGTGGTAT TACAGATTIA TCTATTAAC TCAATCGATGT TTTAACAGGC CTAGACACAG 1140  
 TGAATACTCTG TACAGCTTAT GAATTAGACG GTAAAGAAAT TACTGAGTAC CCAGCAAAC T 1200  
 5 TAGATCAATT AAAACGTTGT AAACCAATCT TTGAAGAGTT ACCAGGTTGG ACAGAAGACG 1260  
 TAACAAATGT GCGTACTTTA GAAGAATTAC CTGAAATGC ACGTAAATAT TTAGAGCGTA 1320  
 TTTCAGAAAT ATGTAATGTA CAAATTCTTA TCTCTCAGT TGGTCCAGAT AGAGAACAAA 1380  
 10 CAAACCTATT AAAAGAAATTG TGGTAGAACT TTATATAAGT CATACACAAT GATTATAAAT 1440  
 ACATAGCCT TCTATCTTTA TTGGTAGGAG GCTTTTGTTA TGCTTGCTTC TGTATCGAIT 1500  
 CGATTATTTA GATAAAAAAT ACTAACGTAA AGGCGATATT TGCTAGTCAT AATTTAGAAG 1560  
 15 YTTAGATGAT AtTtAACGAA AAtTAaGATg aNATAcTGA ATGGTAA 1607

(2) INFORMATION FOR SEQ ID NO: 297:

20 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 3055 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 297:  
 TTAGAAGTAA GCATTTTAAT ATCTTTACCT ACCCATGTAC CAACACCTTC TTTAGGCTTT 60  
 30 GGATTTTCAG CATGGTTATT TGATTTATTC ACCTGTTCCAG ATCTATTAC TTTATTACCT 120  
 TTATTCGGGT TCTCTAATAC ATCAAAATTT AATCTCGGTG AATAAAAAAG ATATATTAAA 180  
 AATCCAAC TA AAAATAATAC ACCGACGACT CTTATAATTA ACTTTTTCAT CAATCAACCA 240  
 35 CCTAAAAAGT ATTAATACTA TTGTAAAAA CAACACATTA ATTAGCAAA TTTCAACACT 300  
 GACATTAAGT TGTCGTTTCG ATAAAACAA ATCATCTTCA GGCATAAATT TAAGTCATA 360  
 40 AAGATTTTCC CTACACTCTA TATCATAACC TATCTTTATG TTTTCAGGTT GAATTTCAIT 420  
 ATTAAGATTA AAATATGTAT AAAAAAATGG ACAGTTAAGG TATCAATGA ATACCATCAA 480  
 CTGTCGACT ATTCTTCATC AAAAAACCTG ATAAACAAA TTGCCTTATC AGATTAGTAT 540  
 45 CATTTGTATA AGCATATTAA TGACCAATG TTGCTTTAAT CAGTAGTCG GTTTCTCCAC 600  
 CTGGATATAA TACATATAAT AATAAGTAAA CTGTACACC TGTAATTGCA GTACAGAACC 660  
 ATATAACTGA AGCGAATGGC CCGAATTTGC GGTGTACATT ATATTATCT TTAATGTCAG 720  
 50 TAATAATTTG AACTAGGCCT AGAATACCAC CAATTGTTGC TAAATTAAAG TGAAAAAATA 780  
 AGAAAAATCGT ATAATATTC TTAATTGATG CTGTCGCCC AAATGCTGTA TTACCGATAA 840

55

|    |   |      |
|----|---|------|
|    | TAACATTTT GTGTTTATTT ATTTCCCTTT TCCAAATAG TCTCCAACCA ATGGCAATTA   | 960  |
|    | AAATGGCACT AATGACAATA CATGTCGTAC TAATCGTTGG TAAATTTGGA ACGCCCATAT | 1020 |
| 5  | TTTTTCATCT AACTTAATTA ATCTAGATCA AAGTAAGTAA TGAACAATC ACAGCTAACA  | 1080 |
|    | CGAAAAAGAT CACTAAATAA TTTAGTGAAT ATATAAACAT TTGTGTTGCC CATTTTGT   | 1140 |
|    | GATCTGAATT TTTCTTAAAT GTTGTTAAAC CTAATGCAAT CCATCTTAA TTTAATAAGG  | 1200 |
| 10 | TTCTCAACCA TaCGAATAG ACACCTAAAT TTATTAGTAA TAAAGTACT GGCAATAAAA   | 1260 |
|    | TAATCAACCA GATAAACATA CTGACACGTG TACGTTTAAA GCCCTTAACT GATGGTAA   | 1320 |
|    | TTGGAATATT TGCAAGTGCA TATTCATCTT TACGTTTAA AGCTAAGGcA TAAAAATGAA  | 1380 |
| 15 | TTGGTGGCCA ACtAAATACA ACTAAAAACA GCGCAATCGC TGTTAAACTA ATTTGTCTT  | 1440 |
|    | CAATTGCAAC CCATCCAATT AGTGGTGGTA CTGCTCCAGG AAAACTCCCA ATCACTGTGT | 1500 |
|    | TCCATGTTGT ATGTCTTTTA GACCATATTG AGTAATAAGA CACATAACCT ACAATCCCCA | 1560 |
| 20 | TAAGACCAAG TACGCTGAT GGTATATTCA ATAAAAACAA ACAATTTCT CCAACTAACA   | 1620 |
|    | TCATACCAAA ACTTAATAGT AATAAATTTT GATCTGTAAT TCTATTATT ACAGTTGGTC  | 1680 |
| 25 | TATTTTGT TT ACTAGGCATA ATACGATCAA TATCTGGTC GTAATAATTA TTTAACGCAC | 1740 |
|    | ATGCGCCACC CATAATTAAA GTAGATCCAA ATAGCATTAA TAAATTTTGA GGTATTGATG | 1800 |
|    | ATAAGAAGGA ATGATTTGTC ATTACAACCTG CTAGCCATGC GCcCGAAAA GCCGGAATTA | 1860 |
| 30 | AGTTACCTTG AACAAAGTCCC ATTTTAAATTA TCTGTTGCAA TTCTTTGAAG TTAAGCTG | 1920 |
|    | TAATATTTTG TGACnAAGTA TGCTCTTGC TCATAATCCC CCTCCTTAAA TTGTTCATA   | 1980 |
|    | TAAGATTATG ATATCTTAGA TTGCATAAAA AGACTAGGTT TAATAAAAT AAATTGTGAC  | 2040 |
| 35 | AAATTAACGA CAAGAGAAAA TGTCAATTTT GTGACACAAA TAACATTTAA TTTATTGCTA | 2100 |
|    | TAATTTATAT GTTAGAAAA TTTAATAAGT AGAATCATGC ATCTAAAAGA GATTAAATAT  | 2160 |
| 40 | TAAGCTTCAA ATTTGAGTAA AGTGGGATTA CATAATTATC CCAATAAAAA AATCATTAAG | 2220 |
|    | ATTAAGTTCT TTTTATGTCG TCCACATACA ATACTGTAA AATTAAATCA TATTTCTG    | 2280 |
|    | GTGGATCCC ATCTTTTCAT ATCCTACAAT CAGGCTCTATT TATAGTATCA TCTCAAAATC | 2340 |
| 45 | GGCTATTAA TCTAATCTC AGTGATGCGT TTTTTATGA TGGGGTGAT AAATTGTTG      | 2400 |
|    | GCAAAAAGAA TTTAAATGG TTAGGTGTCG TAGCAACGTT AATGATGACA TTTGTACAC   | 2460 |
|    | TTGGTGAGC CTTAGTTACC AAAACCGGAT CAGCTGATGG TTGTGGTTCT TCTTGGCCAC  | 2520 |
| 50 | TATGTCATGG TCGGTTGATT CCAGAATTCT TTCCTATTGA TACGATTATT GAGTTAAGTC | 2580 |
|    | ATAGAGCCGT TTCAGCTTG TCTTTATTA TGGCTTATG GTTAGTTATC ACTGCATGGA    | 2640 |

EP 0 786 519 A2

TATTGCAAGC ATTAATCGGA GCTGCTGCTG TTATTGGCA ACAAACGAT TACGTTTTAG 2760  
 cATTGCACCT TGGTATATCA TTAATCAGTT TCTCATCTGT ATTTTAATA ACATTGATTA 2820  
 TTTTCICAT AGATCAAAAA TATGAAGCTG ACGAATTATA TATCAAAAAG CCAATTAGAC 2880  
 GTTTAACATG GTTAATGGCA ATCATCATTT ATTGTGGTGT TTATACTGGT GCACTAGTga 2940  
 GACATGCGGA TGCAAGTTTA GCATATGGTG GTTGGCCATT GCCATTACCA CGATCTTGTA 3000  
 CCACATTCAG GAACAAGATT GGGTCAACT CACGCATCGT ATCAnGGTCn nTTAA 3055

(2) INFORMATION FOR SEQ ID NO: 298:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 748 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 298:

TTCTTATTTA AAGAAGTCAT TTTTAGAAAT TGTTGAGACT TTAAAAAATG ATCCGTATAA 60  
 AATAACACAA TCTTTTGAAA AATTAGAGCC TAAATATTTA GAGCGATATT CAAGAAGAAT 120  
 TAACCATCAG CACAGGGTCG TCTATACCGT AGATGATCGA AATAAAGAAG TATTAATACT 180  
 ATCGGCATGG TCACATTATG ATTAATGAAT ATTCAATATC TGAATAACTT TAATGATAAG 240  
 TTAATTAAAG AAAC TAGTAT TTAAGTGTAG GGAAATAGC GACGTTAATG CGTGTATTAC 300  
 TCTACACTTT TTAATTTTAT AATAGCGCAA GACTAAACAG ATTGAAATTA GTAACAATAA 360  
 AAGAATAACG TATTATAATA AGGAATTTTA AATTGTGACT TTTTCGGAAT ATTAATATTT 420  
 AGAAATATGA GGTITTTAAG CGGATTCCTC ACAAATTTT AAAAATATTT AAGCCTGAAA 480  
 ATGATAAAGC GGTAGGGAAC GTTTTCTGA AAGTTAGTGA TACAATAGTT TTAAGTTGAA 540  
 ATACAGGAGG ATGAATAACA TGAATCAGTC AGTCAAATTA CTTAAACATT TAACAGATGT 600  
 AAACGCCATT GCTGGTTATG AAATGCAAGT TAAAGAAGCA ATGCGTAAct ATATAGAGCC 660  
 TGTcAGTGAT CaaATTATTG AAGATAACTT GGGTGGCATT TTTGGAAAGA AAAATGCTGA 720  
 GAATGGTCAA TACTCAATTA TGGAnTTC 748

(2) INFORMATION FOR SEQ ID NO: 299:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 4718 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear



(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 299:

|    |  |      |
|----|--|------|
|    | ACCTCCGAAT AATTGATTCC ATTAACTTTT TTGTGAAAA ATTAAACAG GCGAAGTCTT    | 60   |
| 5  | CAATAAGTGA AGATGCCAA ATTATAAAAA ATACATTCCA AAAAGAAAAG TTAGGTACAG   | 120  |
|    | TAATTACTAC yCTGGCGCAA GTGGTGGTGT TACGTATAAA CCAATGATGA GTAAAGAAGA  | 180  |
|    | GGCGACTGAA GTTGTTAATG AGGTCATTAC TCTATTAGAA GAGAAAGAAC GTTTGTACc   | 240  |
| 10 | TGGCGGATAT TTATTTTAT CAGATTGGT AGGTAATCCA TCGCTACTAA ACAAGTTGG     | 300  |
|    | TAAGTTAATT GCCAGTATT ACATGGAAGA AAAATTAGAT GCTGTGTGA CCAATTGCSAC   | 360  |
|    | AAAAGGTATT TCATTGGCAA ATGCGGTGTC TAATATTTTA AATTACCAG TAGTAGTGAT   | 420  |
| 15 | TAGAAAAGAC AACAAAGTGA CTGAAGGTTT TACAGTTTCA ATTAATTACG TTTCAGGATC  | 480  |
|    | TTCAAGAAAA ATAGAAACAA TGGTACTTTC GAAGAGAACT TTAGCAGAAA ATTCAAATGT  | 540  |
|    | TTTAGTTGTC GATGATTTTA TGAGGGCTGG TGGCTCTATT AATGGTGTGA TGAATTTAAT  | 600  |
| 20 | GAATGAGTTT AAAGCCCATG TAAAGGGGT ATCAGTACTT GTAGAATCAA AAGAAGTTAA   | 660  |
|    | ACAAGATTG ATTGAAGATT ATACTTCCTT AGTGAAATTA TCTGATGTAG ATGAATATAA   | 720  |
| 25 | TcAAGAGTTT AACGTAGAAC CTGGCAACAG TTTATCTAAG TTTTCATAAA AGGAGTTTTA  | 780  |
|    | GTATTATGAA AATCATTAAC ACAACAAGAT TACCGGAAGC ACTTGGACCA TATTGCGATG  | 840  |
|    | CAACAGTTGT GAATGGTATG GTTTATACCT CTGTCAGAT TCCATTGAAT ATTGATGGAC   | 900  |
| 30 | ATATCGTAAG CGCTGATGTT CAAGCACAGA CAAACAAGT TTAGAAAAA TTAAGGTTG     | 960  |
|    | TTTTGGAAGA AGCAGGATCT GATTGGAATT CTGTCGGA AGCGACCATT TTCATTAAAG    | 1020 |
|    | ATATGAATGA TTTCCAAAAA ATAAATGAAG TGTATGGTCA ATATTTTAAT GAACACAAGC  | 1080 |
| 35 | CAGCGCGTAG TTGTGTAGAG GTTGC CGT TGCCAAAAGA TGTGAAAGTA GAAATTGAAT   | 1140 |
|    | TAGTTAGTAA AATTAAGGAA TTATAATTTT CGATTAAAT GTTTAATCAA GCTTCTAAAT   | 1200 |
| 40 | AAAACAGAGA GATATATACT ATAGGGGGGC TCACTACATG AAAGTGACAG ATGTAAGACT  | 1260 |
|    | TAGAAAAATA CAACAGATG GACGAATGAA AGCACTCGTT TCCATTACAT TAGATGAAGC   | 1320 |
|    | TTTCGTAATT CATGATTTAC GTGTAATTGA AGGAACTCT GCGTTGTTG TTGCAATGCC    | 1380 |
| 45 | AAGTAAACGT ACACCAGATG GTGAATTCCG CGACATCGcg CATCCTATTA ATTCAAGATAT | 1440 |
|    | GAGACAAGAA ATTCAAGATG CAGTGATGAA AGTATATGAT GAAACAGATG AAGTAGTACC  | 1500 |
|    | AGATAAAAAA GCTACATCAG AAGATTGAGA AGAAGCTTAA TCAATTTTAT ATTTAGCGAT  | 1560 |
| 50 | GTAATACATT TGCAATAAGT TGATTGATA CTGTCGATAA AGCATAAAGC TTTGTCGGCA   | 1620 |
|    | GTTTTTTTAG TTGTATTAA TGTTTTTTTA TTTTAAATGA AAGGCTAATA AATATATACG   | 1680 |

EP 0 786 519 A2

|    |  |      |
|----|--|------|
|    | TGATGCTCGT ATTTTGAAG TAAGAAAAA GTTGTTTTTA AAATTACAAC GAATTA AAAA   | 1800 |
|    | CAATGCGCTT TATATGTGA AAGAGTATTG CagATTAAAT TaTAATAATG ACGAaGgTAA   | 1860 |
| 5  | AAATTTAATGG GGGTTAATGT TCATGCGAAG ACACGCGATA ATTTTGGCAG CAGGTAAAGG | 1920 |
|    | CACAAGAATG AAATCTAAAA AGTATAAAGT GCTACACGAG GTTGCTGGGA AACCTATGGT  | 1980 |
|    | CGAACATGTA TTGAAAAGTG TGAAGGGCTC TGGTGTCGAT CAAGTTGTAA CCATCGTAGG  | 2040 |
| 10 | ACATGGTGTG GAAAGTGTAA AAGGACATT AGGCGAGCGT TCITTATACA GTTTTCAAGA   | 2100 |
|    | GGAACAACCT GGTACTGCGC ATGcatGCAA ATGGCGAAAT CACACTTAGA AGACAAGGAA  | 2160 |
| 15 | GGTACGACAA TCGTTGTATG TGGTGACACA CCGCTCATCA CAAAGGAAAC ATTAGTAACA  | 2220 |
|    | TTGATTGCGC ATCAGCAGGA TGCTAATGCT CAAGCAACTG TATTATCTGC ATCGATTCAA  | 2280 |
|    | CAACCATATG GATACGGAAG AATCGTTCGA AATGCGTCAG GTCGTTTAGA ACGCATAGTT  | 2340 |
| 20 | GAAGAGAAAG ATGCAACGCA AGCTGAAAAG GATATTAATG AAATTAGTTC AGGTATTTTT  | 2400 |
|    | CGCTTTAATA ATAAAACGTT GTTTGAAAAA TTAACACAAG TGAAAAATGA TAATGCGCAA  | 2460 |
|    | GGTGAATATT ACCTCCCTGa TGTATTGTCG TTAATTTTAA ATGATGGCGG CATCGTAGAA  | 2520 |
| 25 | GTCTATCGTA CCAATGATGT TGAAGAAATC ATGGGTGTAA ATGATCGTGT AATGCTTAGT  | 2580 |
|    | CAGGCTGAGA AGGCGATGCA ACGTCGTACG AATCATTATC ACATGCTAAA TGGTGTGACA  | 2640 |
|    | ATCATCGATC CTGACGACAC TTATATTGGT CCAGACGTTA CAATTGGTAG TGATACAGTC  | 2700 |
| 30 | ATTGAACCGA GCGTACGAAT TAAATGTCGT ACAGAAATTG GCGAAGATGT TGTATTGGT   | 2760 |
|    | CAGTACTCTG AAATTAACAA TAGTACGATT GAAAATGGTG CATGTATTCA ACAGTCTGTT  | 2820 |
|    | GTTAATGATG CTAGCGTAGG AGCGAATACT AAGGTCGGAC CGTTTGCACA ATTGAGACCA  | 2880 |
| 35 | GGCGCGCAAT TAGGTGCAGA TGTTAAGGTT GGAATTTTTG TAGAAATTA AAAAGCAGAT   | 2940 |
|    | CTTAAGATG GTGCCAAGGT TTCACATTTA AGTTATATTG GCGATGCTGT AATTGGCGAA   | 3000 |
| 40 | CGTACTAATA TTGGTTGCGG AACGATTACA GTTAACTATG ATGGTGAAAA TAAATTTAAA  | 3060 |
|    | AcTATCGTGC GCAAAGATTG ATTGTAGGT TGCAATGTTA ATTAGTAGC ACCGTGAACA    | 3120 |
|    | ATTGGTGATG ATGTATTGGT GGCAGCTGGT TCCACAATCA CAGATGACGT ACCAAATGAC  | 3180 |
| 45 | AGTTTAGCTG TGGCAAGAGC AAGACAAACA ACAAAGAAG GATATAGGAA ATAATCATTT   | 3240 |
|    | ACGTATTTAA AATGGCTAGG ATAAAAGGAT AATCCTATGT AATATTAATG TAATCTTTAT  | 3300 |
|    | GATTTAATGA TTCGCATAGT AATGGAGTTA CATyTTATAT ATAATAGTAA TTGCGTAAGT  | 3360 |
| 50 | AAATAATTGG AGGACTATAA ATGTTAAATA ATGAATATAA GAATTCGTCA TTAAGATTG   | 3420 |
|    | TTTCATTGAA AGGAAACGAA GCATTAGCGC AAGAAGTTGC TGACCAAGTA GGAATTGAAC  | 3480 |
| 55 |  |      |

EP 0 786 519 A2

|    |  |      |
|----|--|------|
|    | GTATTCGTGG TTGTGACGTA TTTATTATTC AACCAACATC ATATCCTGTG AATCTACATT  | 3600 |
|    | TAATGGAAIT ATTAATTATG ATTGATGCTT GTAAACGTGC TTCTGCAGCA ACAATCAATA  | 3660 |
| 5  | TTGTAGTGCC ATATTATGGA TATGCAAGAC AAGATAGAAA AGCCCGTAGC CGTGAGCCAA  | 3720 |
|    | TCACCTGCTAA ATTAGTTGCA AACTTAATCG AAACAGctGG CGCAACTCGT ATGATTGCGT | 3780 |
| 10 | TAGACTTACA TGCACCACAA ATTCAAGGAT TCTTTGATAT TCCAATTGAC CACTTAATGG  | 3840 |
|    | GTGTGCCAAT TCTTGTCTAAA CATTTCAAAG ATGATCCGAA TATTAACCCA GAAGAATGTG | 3900 |
|    | TGTTTGTTC CAGACCCATG GCGGshTTAC ACGTGCACGT AAATTAGCTG ACATTTTAAA   | 3960 |
| 15 | AACTCCAAT GCAATTATAG ATAAACGTGC TCCTAGACCA AATGTTGCTG AAGTGATGAA   | 4020 |
|    | CATTGTGTGT GAGATTGAAG GACGTACGGC AATTATTATT GACGATATTA TTGATACAGC  | 4080 |
|    | AGGTACAATC ACTTTAGCTG CACAAGCATT AAAAGATAAA GGTGCTAAAG AAGTATATGC  | 4140 |
| 20 | TTGTTGTACA CACCCTGTTT TATCAGGACC GGCTAAAGAA CGTATCGAAA ATTCTGCTAT  | 4200 |
|    | AAAAGAATTA ATCGTAACAA ACTCAATTCA TTTAGATGAA GATCGCAAA CACTTAACAC   | 4260 |
|    | TAAAGAATTA TCTGTTGCTG GTTTAATCGC ACAAGCTATC ATTGCTGTAT ACGAAAGAGA  | 4320 |
| 25 | ATCAGTTAGC GTATTATTG ACTAATATTT AAAAGGCGTT TGACGAACAT ATTCCAAACG   | 4380 |
|    | TGTATAATAG TTGTTGCTG GATTATACGA ATAAATAAAC ACTTGCAAGC AACGATGATG   | 4440 |
|    | TTGATGGGTA AGTGAGGTGC TCGTTTGTAG CAAAAATGAA AGGTGGAAAT GAGAATGGCT  | 4500 |
| 30 | TCATTAAAGT CAATCATCCG TCAAGGTAAA CAAACACGTT CAGATCTTAA ACAATTAAGA  | 4560 |
|    | AAATCTGGTA AAGTACCAGC AGTAGTATAC GGTACGGTA CTAAAAACGT GTCAGTTAAA   | 4620 |
|    | GTTGATGAAG TAGAATTCAT CAAAGTTATC CGTGAAGTAG GTCGTAACGG TGTATCGAA   | 4680 |
| 35 | TTAGGCGTTG GTTCTAAAAC TATCAAAGTT ATGGTTGC                          | 4718 |

(2) INFORMATION FOR SEQ ID NO: 300:

|    |                               |
|----|-------------------------------|
| 40 | (1) SEQUENCE CHARACTERISTICS: |
|    | (A) LENGTH: 3181 base pairs   |
|    | (B) TYPE: nucleic acid        |
|    | (C) STRANDEDNESS: double      |
|    | (D) TOPOLOGY: linear          |

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 300:

|    |   |     |
|----|---|-----|
| 50 | AAATGCATA TAAATACATA TTAAGGAGGA TTTTATGAAT TTTCTTAAAC CTGCAAAGCA  | 60  |
|    | TATTAAGCCT TTGCGAGAAA ATCAGATAGA TGATACCTAT AAACGATTAC GTCTCCAAGT | 120 |
|    | ATTCTCTGGT ATTTTCATCG GTTACGCTGG GTACTATTTA TTACGTAAAA ACTTTTCGTT | 180 |

55

EP 0 786 519 A2

TGCTGTTTCC ATCGCATATG GATTAGTAA GTTCTTTATG GGTACTGTAA gTGATCGGAG 300  
 CAATGCTCGG ATATTCTTAG TTCTTGGATT AGCACTCACT GCTATCGTCA ATTTGTTAAT 360  
 5 GGGATTTGTG CCGTTCCTTA CATCAGGTAT CGGTATTATG TTTGTCTAT TATTCTTAAA 420  
 TGGATGGTTT CAAGGTATGG GCTGGCCACC TTCAGGCCGT GTTCTCGTTC ACTGGTTTAG 480  
 TGTAAAGTGA CGGGGAAGTA AGACTGCCCT TTGGAACGTT GCGCATAAATG TTGGTGGAGG 540  
 10 TATTATGGCA CCTATTGCTG CTGGGGTAT TACAACAAC TGCATTATCA ACTTTGGTTA 600  
 TTTAAAAGGT TTGGAAGGTG TATTCAATTA CCCTGCACCT TTAGCACTTA TCATTGCCGC 660  
 AATTTCATAC GTATTGATTA GAGACACACC TCAATCTCAA GGTTCACCTC CAATCGAAAT 720  
 15 TTATAAAAT GACTTTGCTA CAAGCGATAA GAAAACATTA GAAACAGAAT TAACACAAA 780  
 AGAAATTTTA TTTAAATATG TACTGAACAA TAAATGGGTA TGGCAAAATG CCTTTGCAA 840  
 TATATTGTTT TATTTCGTGC GTTATGGTGT ACTGATTGG GCGCCAGTCT ACTTAAGTGA 900  
 20 AGAAAAACAT TTCGACTTAA AAGCATCAGG TTGGGCATAC TTCTTATACG AATGGGCTGG 960  
 AATTCTGGT ACATTATTAT GTGGTTACAT TTCTGATAAA TTATTCAAAG GTCTCGTGG 1020  
 25 ACCTGCAGGT TTCTTCCTTA TGTTAGGTGT CACAGTATT GTATTAAATTT ATTGGTTAAA 1080  
 TCCTCCAGSC AATGCTTGGT TAGACAATGT CTCATTAAIT GCCATTGGTT TCTTAATATA 1140  
 TGGACCAAGT ATGTTAATTG GTTTACAAGC ATTAGATTAT GTACCTAAAA AAGCAGCTGG 1200  
 30 CACAGCAGCT GGATTAACAG GATTATTGGG TTATCTGTIT GGTGCTGTAA TGCCCAACAT 1260  
 CGTCTTAGGT GCTGTAGTTG ATAAATTCGG ATGGGATGTC GGTTTTATTT TATTAAACGC 1320  
 AATTAGTGTG TTTGCAATGT TGAGCTTTAT CCTCACTTGG AATAAGTAG GACAAGAAAC 1380  
 35 CGTTCATCAT TAAATGATAA AAAATAAAGT CATATGTTA TCTTATCGAA AGATGATATA 1440  
 TTCATCTCTT ATAAGTTCAA CCATATGACT TTTTATTAGT ATTCAAAAAA ATATTACAT 1500  
 TGCCACTTTT GTGTTTGGCC TGCTGTTTTA TTCAATTGAT TACACCACCT AGGATAAACT 1560  
 40 CTAAAGGCCA TTTTCCCTTG ATACTTAGAT GAAGCTAAGA TACCTTTTTC TACCAATAAT 1620  
 TCCCTAGGAA ATAGAAAGTA GCCATTTAAT TCATCATCGA TAACAGCAAC AATCAGGTAA 1680  
 45 TCACGAAACT CTTCAAATTT GTATGGCCGA TTATAATTAT CTTCTGCTTT TGTCACAACT 1740  
 GTCACGAAAT ATCCCGATTT AGTCGGTGTI TTCTTAGCTA ACCTACTTTG ATATGTTTCC 1800  
 TCTTTAAAGC TAAATGTTAA TGCTCGTAA TCTTGATTAT ATTTTCTTC AGTTAAATCT 1860  
 50 TTAACCTCTG ATTCTTCACT AAAAATATTT TTCAACAGTA TTTTAGATT ACACATATGC 1920  
 CAATCTCACT TTATTTTTTC ATAATCGTAT CATATATTTA TTTTTTCGA AAAATACACT 1980

55

EP 0 786 519 A2

|    |   |      |
|----|---|------|
|    | TGTATCAACT GTCAAACGAT CACGTAATAA ATAGACGATT AACATTGcTG CTAAAGCACA   | 2100 |
|    | AAGTGACTCA GCAATTAATA ATGACCAAAAT GACACCTGTT AGTCCAAACA AAGCATTCTAT | 2160 |
| 5  | AATAAATAAT ACTGGAATGA TAATTGCACC TTGTAAATG GCCATAATTG TAGCACCACG    | 2220 |
|    | ACCTTGCCCA GTGCTTGAA GCATACCAGT AAACAAGAAA CCTATACCAT TTAATAATAA    | 2280 |
| 10 | TGATGCCATT GTTACTTTCA AAATAAATGc CGCCATCTCA ACAATGGCTT GATCAGTAGT   | 2340 |
|    | AAATAGTCCG ACCATATGAT GTCCAATTGT AAATACAGCA CTCATACATA CAACAAGAT    | 2400 |
|    | AACGCCGATA GACATGATAA CTGCTTTGAT AACGTCTTTC ATACGGCCCTT TATTTGCCAT  | 2460 |
| 15 | AAAGTTATAT GCAATTAGTG GTACAACACC TTCACATAAT CCCATGATAA TAAGTTCTGG   | 2520 |
|    | aAATTGCACA AGTCTAAATG AGATACCATA ACTTGCaATC GCGAAGTTTC CATAAATGTGc  | 2580 |
|    | TAAAAATAAA TTTAAACTA ATCCTGTGAA TCCCATTAAG ATACTCATTa AAAATGCAGG    | 2640 |
| 20 | AATACCGATT TTAAGATTt CAGAAAGCAT TTCTTTATTA GGTTCGCAA GTTTAATATt     | 2700 |
|    | AAC TGACACA ACGTCACTAT TTTTCATAAA ATAAATGATA AAGAACAGAG CAGCAGCAAC  | 2760 |
|    | ATTACTGATT GCAGTACCCA AAGCTGcACC AACACGTTT aAATCAAAAC CAAAAATTAA    | 2820 |
| 25 | AATTGGATCT AAAATAATAT TTAAGCCTAC ACTAGCTAAC ATACCAATCA TAGAAACCAT   | 2880 |
|    | TGGTGCCCCA ATTGCAGTG CAATTTGTTT TAATATGAAG AACAAAAATa CAAAAGGTGC    | 2940 |
|    | ACTTAAAAAC ATTACTTTCA AATAATTACT TGTtAAAGCT AACGTTTCAC CTCTCGCCCC   | 3000 |
| 30 | TAAAATTGCT GCGATTTGAT CACTGAATGG TAAAGTAACT AAAATCAGa TAAGTCTTAG    | 3060 |
|    | TGCAATACCA CCATAAATAG AGAAACTACT TACAAATTa CTCTAmTAT AGTCTTTGCG     | 3120 |
|    | ACCTAATAAA CGTGAAATAT AAGTCTCTGC ACCAAGGCCA AATAAATTAC CTAACCCCAT   | 3180 |
| 35 | T   | 3181 |

(2) INFORMATION FOR SEQ ID NO: 301:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4029 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 301:

|    |  |     |
|----|--|-----|
|    | TCATCACCTT CAAATAAATA GGCATGCGAT AATTtATTG AATGATATGC ATTcGTCAAT | 60  |
| 50 | TGTtGTCTGT CATCCATTCA TTACAAACTC cTTTTGCTT TTATATAAAA AGGTTGCCAA | 120 |
|    | AGAGCAACTA ATTACTGCAA CAATAGCGAC CATGATTtGT GATTATTtTT AGAACCAAA | 180 |

EP 0 786 519 A2

|    |  |      |
|----|--|------|
|    | TTATTAATAC ACTATTTTAA GTATCTCTAA TTGATGATTG TATTATAGAA TTAAAAATTGA | 300  |
|    | TGGAATGCAT CAACTGGCAT AACAAATACA GTAGCACCGC CAACCTCAAC TTCAACTGGA  | 360  |
| 5  | TATGGAATGT ACGAATCCGC ACTGCCTCCC ATAGGTGTAA TAGGTGAAAC CAACTGTTCT  | 420  |
|    | CTATTACCAC ACGTTTGATT AATCACAGAC AATATTTCAT CTACACGGTC ATCATTGACA  | 480  |
|    | CCACATAAGA ATGTTGTATT ACCCGCTCTT AAAAACCCAC CTGTTGTGTG CAATTTGTIT  | 540  |
| 10 | GCTCTAAAGT TATTTTAAAC AAGTTGATCT GCAAGTTCCT GACTATCTTG ATCTGTGACG  | 600  |
|    | ATCGCTATAA TCATTTTCAT TTATAACAC CTCTTCTAAC AATTATATCA TATCTTTTCT   | 660  |
| 15 | AAATATTGA TGATAGTTG ATACGTGTCT TCAACAACAT TTTCAGAGG TTGATCTGCA     | 720  |
|    | TTAACGCTTT TGAACGGTGG TGATTCAITA TGAATGATTT CTGGTAACCT TTCAATTACT  | 780  |
|    | TTTTGTGAA CTTTAAATCT TCTTGATCTA ATCTATTTTG ATCTCTTGAA TTTTAAATAA   | 840  |
| 20 | TACGTTGGCG ACCTACTTCA GCACTAACAT TTAATAAAT CGTCAAGTCT GGATATAATC   | 900  |
|    | CATTTATTGC AAATTCGTTT AATGCTCTTA CTCTTCAAC GCCAATCCCT CTAGCATAAC   | 960  |
|    | CTTGATAAGC TAATGAACATA TCGATATAGC GATCACACAA CACAACCTTA CCTTCTTTTA | 1020 |
| 25 | AAGCTGGTAT GACCTTTAAT ACAAGATGTT CTCTTCTAGA TGCAGCAAAAT AACATTGCTT | 1080 |
|    | CAGTCTAAT GTCCATATCA TTGCCCTTCT AATACAATTT TACGTATTTG TTCACCACTA   | 1140 |
|    | GGAAACACCAC CTGGTTCCTT AGTCATAATG ACATCATAAT CTTTACTTAA TCTATGGTAA | 1200 |
| 30 | ACTTCATTAA TTACAGTTGT TTTTCCAGAG CCTTCTGGGC CCTCAAAAGT TATAAAAGCT  | 1260 |
|    | GACATTTTAT TCATCTCTAA CTAAAATTTT ATTATTTTTA ATTCCTTCAA CTATCATTC   | 1320 |
|    | AGTTTCCAGA TATTCATTTA CCAATTCTAT CATATTTTCA GTTATTGTTT CTCCTTTGAA  | 1380 |
| 35 | AATAATAGGA ATCCCTGGCG GATACGGGAC AATATGTCGC GCCAGAACTT TACCTTTGCC  | 1440 |
|    | TTTTGAAATC ACACCAAGTA ACATATTCAA AGCGTTTGG TTTATAGTTA CCTTCAGTgG   | 1500 |
| 40 | TTAAAAGTTG TGTTTGTTTA ACTTTAGAAA CTGATTTTTT CGGTAAAATC ATATCTTCAA  | 1560 |
|    | TTTACGTAA AAGCGAATCA AATAAATACG TATCATCATG ATGCCATAAC GGCAATATCG   | 1620 |
|    | CTAATGCTTG ATAGTGCTCC GCTAATTTCTA AATAGATATG TGCATTCATA AACCAATTTT | 1680 |
| 45 | GAATATCATG ACCTGTAAAA CCTTCATATT TTATCAGCAA CTTCACCGGA TCATCAACTT  | 1740 |
|    | GAGCATTTTC AAAACCCCTC TTCTCCAAAC ATTGATTAAT TTGCGCTCTC TTATCAAAAA  | 1800 |
|    | ACACGGTACT ATCATATGTT TTATAGAACT CGGCAGCTGA CTCTAAACTA GCCATTAATCA | 1860 |
| 50 | AATACGAAGG ACTAGATGTT TGGAAAGTAGC TTAGATATTC TATAATAGIT TCTCTATAAG | 1920 |
|    | GTGCATTTTT ATGAATATAA AGTACCGAGC CCATCGTTAA AGCTGGTAAC GTTTTATGAA  | 1980 |

55

|    |            |             |             |            |             |             |      |
|----|------------|-------------|-------------|------------|-------------|-------------|------|
|    | CAAAGTGGCG | GCCGTGTGCT  | TGTCATGA    | GTACAGGAAT | ATTTAATTGG  | TGCAAAGATT  | 2100 |
|    | TGATAACCTC | TTCATACATTA | AATGTTTCAC  | CGTAATAGTT | AGGATAAGTC  | AACACAGCAA  | 2160 |
| 5  | GTTTGTGACC | GTCAATTATTC | AAACGGCTTA  | AATTAACTTT | ATTATAATGA  | TTCGTTAAACG | 2220 |
|    | GACTTTGATG | CGTTTCAATA  | AAATGCCCTT  | CTTGTGGCT  | AATATCGAGC  | GCAITGAACA  | 2280 |
|    | CAGATTATG  | TACATTTCCT  | GCCATTAAGA  | TATCGCCTTT | TTTCTGTGAA  | AAAGACTGGA  | 2340 |
| 10 | TGACAGATAA | TATTCCTGAA  | GTGGTGCCAT  | TCACTAAGAA | ATAAGCATCA  | TAATCTGAAT  | 2400 |
|    | GTTTCTCCAC | CTGCTTCATA  | CTTTCGAAAA  | TGACTTCTTC | AGGATGATGT  | AAATCAACTn  | 2460 |
|    | AATCCAGGTA | TTTCAGTTTT  | ATCCATTGTC  | ATTGATAATT | GAGATAAATG  | ACCGATAGTC  | 2520 |
| 15 | ATATTTTTAT | GACCCGGAAC  | ATGCAAAGAA  | ATCGCTTCTT | CTTGATTATA  | ACTTTCTAAT  | 2580 |
|    | TTATTTAAAA | TAGGTGCTT   | CATGATATAC  | GCTTCCTTTA | TTTACACTGT  | TTTGGAAATTA | 2640 |
| 20 | GTTACTTTCA | AAAGTATTAA  | TTATATAGTA  | ACACTTCTTT | GACAAAAGTT  | AGTGTACTTT  | 2700 |
|    | ATGCAATAGC | TTGTCTATTG  | TATAATAATT  | AATTTCCTTT | TTGTACTTCG  | ATTTAAAAAGA | 2760 |
|    | TATTAGACAT | AAAATCTAAA  | AACAGCAGTA  | AGATGATTTA | TGATTAAAAA  | CTATCTTACT  | 2820 |
| 25 | GCTGTTCACT | TTTTATAATA  | CTTCTGAATG  | TCTTCACTTA | TACTTCTAGT  | CACAGATTTA  | 2880 |
|    | AATAATCAAA | AGTGACATT   | ATTAANAATAT | CAATTCACA  | CTCAATGCGG  | CTCATCGCAT  | 2940 |
|    | TCATTTCTTG | TCTAGCAACG  | TTCTACTCTA  | GCGGAACGTA | AGTTAGCTAC  | CATCCTCGCT  | 3000 |
| 30 | AAGAACCTTT | CTTGACTTGT  | GACAATCGCT  | TGCTTCTTTC | CTCTCCTTCG  | GCTCTCGCTT  | 3060 |
|    | ACTCATTTAG | CTCTACTAAA  | CTCGTTGCGC  | TCCTTCTCG  | TTTCGTCAGA  | TTCAACGTT   | 3120 |
|    | TTCACTTCGC | CAAGCCATTT  | TTCTTTGTGT  | TTACTTTTTA | TTTTGACGTT  | TTAGACATAA  | 3180 |
| 35 | AAAAAGAC   | CTACGGTCT   | CAACTTGCCT  | GGCAACGTTT | TACTCTAGCG  | GAACGTAAGT  | 3240 |
|    | TGGCTACCAT | CGTCGTAAA   | GACCTTTCTT  | GACTTGTGAC | AATCGCTTGC  | TTCTTCTCTC  | 3300 |
| 40 | TCCTTCGGCT | CTCGCTTACT  | CATTAGCTTC  | TACTAAACTC | GTTCGCGTCT  | TTTCTCGTIT  | 3360 |
|    | CGTCAGATT  | AAACGTTTTT  | ACTTGGCCAA  | GCCATTTTTT | TTTGTGTTTA  | CTTTTTATTT  | 3420 |
|    | TGACGTTTTA | GTACATAAAA  | AAAGAGACCT  | TGCGGTCTCA | ATGCGGCTCA  | TGCGATCCAT  | 3480 |
| 45 | TTTTTGGCTG | GCAACGTTCT  | ACTCTAGCGG  | AACGTAAGTT | GGCTACCATC  | GwCGCTAAGa  | 3540 |
|    | acCTTCTCTG | ACTTGTGACA  | ATCGCTTGCT  | TCTTCTCTCT | YCTTCGGCTC  | TGCTTACTC   | 3600 |
|    | ATTTAGCTCT | ACTAAACTCG  | TTGCGCTCTT  | TTCTCGTTTC | GTACAGATTCA | AACGTTTTCA  | 3660 |
| 50 | CTTCGCCAAG | CATTTTTTCT  | TTGTGTTTGC  | TTTTTATTTT | GACGTTTTAG  | ACATAAAAAA  | 3720 |
|    | AAGAGACCTT | GCGGTCTCAA  | TGCGGCTCAT  | CGCATCCATT | TTTTGCTGCT  | CAACGTTCTA  | 3780 |

55

CATGGGAACA GGTGTGACCT CCTGTCTATA GTCACCAGAC ATATGAATGT AATTITATACA 3900  
 TTCAAACTA GATAGTAAAT GAAAAGTGA TTTGTCTCG CAAAACATTT ATTTTGGATT 3960  
 AAGTCTTCGA TCGATTAGTG ATTCGTGCAG CTCACATGT GCACCATGTG TGCCACCTCG 4020  
 GAACCTATT 4029

(2) INFORMATION FOR SEQ ID NO: 302:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 7159 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 302:

GATGCAATAG TTGAGTAGTT ATAAGTAGCT ACATACAACC GCTCAAATAT AGGTTCAAGA 60  
 ACATTTTAA TGAAGAAAGC GATTTTCCCG CAGAGTGCAA ACGCTTGACT CAATAACGAA 120  
 ATGTCACCTA TGTATGGCGT GACTTATTAT ACAGGAGGTG CAAAGTATGT TTGCTATTAT 180  
 TGAACACAGT GGAACACAAA TCAAGTAGA AGAAGGTCAA GAAATCTTCG TTGAAAAATT 240  
 AGACGTAAC GAAGGAGATA CTTTTACATT TGATAAAGTA TTATTTGTAG GTGGAGATTC 300  
 AGTTAAAGTT GGAGCGCCAA CAGTTGAAGG TGCAACAGTT ACTGCTACTG TTAATAAACA 360  
 AGGTGCGCGT AAAAAAATCA CTGTATTAC ATACAAACGT CGTAAAAATT CAAAACGTAA 420  
 AAAAGGCCAT CGTCAACCAT ACACTAAATT AACATCGAT AAAATCAACG CGTAATTATT 480  
 ATGATTACTG TTGATATTAC AGTTAATGAT GAAGCAAAG TAACAGACGT TATTATGGAT 540  
 GGCCATGCTG ACCATGGTGA ATATGGTCAT GATATCGTTT GTGCTGGAGC TTCAGCTGTA 600  
 TTGTTTGTTA GTGTTAATGC GATTATAGGA TTGACATCTG AGAGACCAGA TATCAATTAT 660  
 GACGACAATG GTGGTCATTT TCATATAAGA AGCGTTGATA CAAACAACGA TGAAGCGCAA 720  
 CTAATCTCTC AAACAATGCT TGTGTCTTTA CAACTATTG AAGAAGAATA TAATGAGAAT 780  
 ATTAGATTAA ATTATAAGTG AGGTGCATTC CGATGTTAAA ATTAACTTA CAATCTCTCG 840  
 CATCTAAAAA AGGGGTAAGT TCTACAAAAA ACGGACGTGA CTCTGAATCA AAACGCTTAG 900  
 GTGCTAAACG TGCTGACGGT CAATTCGTAA CAGGTGGTTC AATTTTATAT CGCCAAACGT 960  
 GTACTAAAAA TTACCTGGT GAAAATGTAG GTGCTGGTGG CGATGATACA TTATTCGCTA 1020  
 AAATCGACGG CGTTGTTAAA TTCGAACGTA AAGGTGCGCA CAAAAACAA GTTTCTGTAT 1080  
 ATGCGATAGC TGAATAATTT TGTCTAGTTA ACACCAAGAG TGAATCTTCT GGTGTTCTTT 1140





|    |   |      |
|----|---|------|
|    | TGTTGGTATG TTGGCACGTG TACTAGATGT TATATCAAAG TTAGAACTAT CTGTATTAAC | 3060 |
|    | GATTTCATCA AGTATTCCAA TGGAAAGAAA AGCAACAATA ACATTATCAC TGAATGCTAA | 3120 |
| 5  | ATCTAAAGAA ACTTCAGTAG AAGATGTTAT TGGCGCTTG AGAAATTTAG ATTATGTATC  | 3180 |
|    | AAAAGTAGAA TTAATTAGTA TGAGTATGTA AGGATGTGCC TATAATGTAC GCGTATGTCA | 3240 |
|    | AAGGTAAGTT AACACATTTA TATCCTACAC ACGTAGTTGT TGAAACTGCT GGTGTTGGTT | 3300 |
| 10 | ATGAAATTCA AACACCAAAT TCTTATCGTT TTCAAAAGCA TCTAGATCAT GAAGTTTAA  | 3360 |
|    | TTCATACATC TTTAATGTGT CGTGAAGATG CACAATTATT GTATGGATTT AGTAGTGAAG | 3420 |
| 15 | AAGAGAAAGA TATGTTCTTG AGTTTAATTA AAGTTACTGG TATTGGTCCG AAATCAGCTT | 3480 |
|    | TAGCTATTTT AGCGACAAGT ACGCCTAATG AAGTAAACG TGCCATTGAA AATGAAAATG  | 3540 |
|    | ATACGTATTT AACTAAATTC CCAGGAATTG GTAAGAAAAC GGCAAGACAG ATTGTCTTAG | 3600 |
| 20 | ATTTAAAGG TAAAGTGAAA ATTACTGAAG AAGATAGCGA TTCATTATTA CAAGTAGACG  | 3660 |
|    | CTACTTCGAC GGTGCAAGAT CAATTCTGTC AAGAAGCAAT GTTAGCGTTA GAAGCATTAG | 3720 |
|    | GTTATTCTAA ACGAGAGCTT GCAAAAGTTG AGAAAACGTT AAATAAAAAA AAATAGACT  | 3780 |
| 25 | CAGTTGATGA AGCTGTTAAG GCAGGTCTTC AATTAGTTGT ATCTTAATTT TAAATAGATT | 3840 |
|    | AATAGGGGAA GTGTTGTCAT GAATGAGCGT ATGGTTGATC AATCAATGCA TAGTGAAGAA | 3900 |
|    | ACTGATTTTC AATTGTCGCT TAGACCTACG AGATTACGAC AATATAATTG TCAAAATTCA | 3960 |
| 30 | ATAAAAAGTA ATTTAGAAAT ATTTATTAAA GCGGCTAAAC TTCGTCATGA ACCATTAGAT | 4020 |
|    | CATGTATTGC TTTTGGCCCC CCTTGATTIA GGTAAAGCAA CATTATCTAA TATCATTGCC | 4080 |
|    | AATGAAATGG AAGTTAATAT ACGTACAGTA TCAGGGCCTT CATTAGAAAG ACCTGGTGAT | 4140 |
| 35 | TTGGCTGCAA TTTTATCAGG ACTTCAACCT GGAGATGTTT TGTTTATTGA TGAATACAC  | 4200 |
|    | AGACTGAGTA GTGTTGTTGA AGAAGTGTTA TACCTGCAA TGAAGATTIT CTTTTTAGAT  | 4260 |
| 40 | ATTATCATTT GTAAAGGCGA TGAGGCTAGA AGTATCCGTA TCGACTTACC TCCATTCACT | 4320 |
|    | TTGGTAGGTG CAACAACGCG AGCTGGCAGC TTAACAGGTC CACTAAGGGA TCGATTGGT  | 4380 |
|    | GTGCACTTAA GATTAGAATA TTATAACGAA TCAGATTTAA AAGAAATCAT TATTAGAACA | 4440 |
| 45 | GCTGAGGTTT TAGGCACAGG TATTGATGAA GAAAGTGCCA TTGAACCTTG TAAAGCTTCT | 4500 |
|    | AGAGGGACTC CAAGAGTAGC AAATCGACTA TTGAAGCGGG TAAGAGACTT CCAGCAAGTG | 4560 |
|    | AATGAAGATG AACAAATATA CATTGAAACA ACGAAGCAGC CATTAGGTTT ACTTCAAGTT | 4620 |
| 50 | GATCAACACG GACTAGATTA CATTGATCAT AAAATGATGA ACTGTATTAT TAAGCAGTAT | 4680 |
|    | AATGCGGCAC CTGTGTTT AGATACGATT GCCGTACAA TTGGTGAAGA ACGTATTACA    | 4740 |

55



|    |   |      |
|----|---|------|
|    | AAAAATTAAA AGCATCACAC CTAAGTAAAC AGTTCaATCA TCTTAAAAAA TCCTGGGACA   | 900  |
|    | CTTCATACCT GTCTCAGGAT TTTTThACAA ATTGAATCAG CCTCATAACA TTAaATTATT   | 960  |
| 5  | TTATCGTACA TTAaATTAA TAATAACAAC TGATTTTAT AAGAATAAAG TATCGAmCCA     | 1020 |
|    | TAGTAGATAC ACAATAATA CAAATGAAAC AATTTAACCT GAAAGCTTaa ATAAATATTA    | 1080 |
|    | TCAAGTTAAT AAACAATTAA TTTTtagATG GATTCACTAA AAATCGTAAA AAAGCACAAT   | 1140 |
| 10 | TTGTATTTTA CAAACATTAA TTAaaaaAGA AAGCAAGACA TTCGTGCAAT CGGTTACCTT   | 1200 |
|    | AAATTGTTTA CAACTGTCAA CAATACCAAG GTTTTATTAA CTATATTCT CACAAAAATTA   | 1260 |
| 15 | GCTTTTAGCA TTCCAAACAA AAAAGGTTAA ATCGAACGGA ATTATGGCAT TTTTAACTTA   | 1320 |
|    | ATTGTAaaaa AAGTTGATAA TGGTCAATTG TTAATGAACA GTTAATTATA ATAACGCCCA   | 1380 |
|    | AAATATATTA TTATTTAATT AAGTTAAATA AAATTATAGA AAGAAAGTGA AACTTATGCT   | 1440 |
| 20 | TAAAAATAAA ATATTAACCT CAACCTTATC TGTGAGCTTA CTGCCCCCTC TTGCCAATCC   | 1500 |
|    | GTTATTAGAA AATGctAAAG CTGCTAACGA TACTGAAGAC ATCGGTAAAG GAAGCGATAT   | 1560 |
|    | AGAAATTATC AAAAGGACAG AAGATAAAAC AAGTAATAAA TGGGGCGTGA CTCAAAATAT   | 1620 |
| 25 | TCAATTTGAT TTTGTAAAGG ATAAAAAATA TAACAAAGAT GCTTTGATAT TAAAGATGCA   | 1680 |
|    | AGGATTCATT AGCTCTAGAA CAACATATTA CAACTATAAA AAAACTAATC ATGTTAAAGC   | 1740 |
| 30 | TATGCGATGG CCATTCCAAT aTaatATTGG TTTAAAAACA AATGATAAAT ATGTTTCITT   | 1800 |
|    | AATTAATTAT TTACCTAAAA ATAAAAATGA ATCTACAAAC GTGAGTCAGA CATTAGGATA   | 1860 |
|    | CAATATCGGT GGTAAATTTCC AATCAGCCCC ATCACTCGGT GGTAAATGGAT CATTTAACTA | 1920 |
| 35 | TTCTAAATCG ATTAGCTATA CACAACAAAA TTATGTAAGT GAAGTAGAAC AACAAAACTC   | 1980 |
|    | AAAAAGTGTT TTATGGGGCG TCAAGCGGAA TTCACTCGCC ACTGAATCAG GTCAAAAATC   | 2040 |
|    | AGCCTTTGAT AGCGATTTAT TTGTAGGCTA CAAACCTCAT AGTAAAGATC CTAGAGATTA   | 2100 |
| 40 | TTTCGTTCCA GACAGTGAGT TACCACCTCT TGTACAAAGT GGATTTAAAC CTTCATTAT    | 2160 |
|    | CGCCACAGTA TCTCATGAAA AAGGTTCAAG CGATACAAGC GAATTTGAAA TTACTTACGG   | 2220 |
|    | AAGAAACATG GATGTCACCT ATGCCATTAA AAGATCAAGC CATTATGGCA ACAGTTATTT   | 2280 |
| 45 | AGACGGACAT AGAGTCCATA ATGCATTTGT AAATAGAAAC TATACTGTGA AATACGAGGT   | 2340 |
|    | CAATTGGAAg ACTCATGAAA TCAAGGTGAA AGGACAGAA TGAATGAAA ATGAATAAAT     | 2400 |
| 50 | TAGTCAAATC ATCCGTTGCT ACATCTATGG CATTATTAT ACTTCTCGGT ACTGCTAATG    | 2460 |
|    | CTGAAGGTAA AATAACACCA GTCAGCGTAA AAAAGTCTGA TGACAAAGTT ACTTTATACA   | 2520 |
| 55 | AAACAACAGC CACAGCAGAT TCTGATAAAT TTAaaATTTC ACAGATTTTA ACATTTAATT   | 2580 |

ACTCAGGCTT TGTGAACTT AATCCTAATG ACTATGACTT TTCAAAATTA TATTGGGGAG 2700  
 CTAATATACAA TGTATCTATA AGCTCACAAT CTAATGATTC AGTAAACGTC GTTGATTATG 2760  
 5 CACCAAAAAA TCAAAATGAA GAGTTTCAAG TTCAAAATAC TTTAGGCTAT ACATTGGGTG 2820  
 GTGACATTAG TATCTCTAAT GGTTTATCTG GTGGACTTAA TGGAAATACA GCTTTTTCTG 2880  
 AAACAATTAA TTATAAACAA GAAAGTTACA GAACAACATT AAGTCGCAAC ACAAAATTATA 2940  
 10 AAAATGTTGG CTGGGGAGTT GAAGCACATA AAATTATGAA TAATGGTTGG GGACCTTATG 3000  
 GAAGAGATAG CTTCACCCCA ACATATGGTA ATGAACTCTT CTAGCTGGC AGACAAAGCA 3060  
 GTGCATACGC TGGCCAAAAC TTCATAGCGC AACACCAAAT GCCATTATTA TCTAGAAGTA 3120  
 15 ACTTCAATCC AGAATTTTTA AGCGTACTAT CACACAGAC 3159

(2) INFORMATION FOR SEQ ID NO: 304:

20 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 3821 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 25 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 304:

30 GCAAAAACTT TTCTCCAAC TATGTCAGACT ATAAATGGCAT GAACAAAATT GATATGATGA 60  
 ATCAAAATTAA AGTTGATACG ATGTTGCATG GTTATCACGC TGGATTTTAA TTTGACTTAT 120  
 TGATTACAGT TGTAGTTTC TTTTGTTCAT TTATGTITACA AGTAAGAAA AAAGAAGTTG 180  
 35 ATTCTCGTCA ATAAATATAA GTTGCTAGCT ATATAAAGCT TTTTAGCAAA AGTTCAACAT 240  
 ATTGACTTAT CCGGCATTTC AGATTAAAT ATTTTATTTC CGATTAGAAT AATAAGAATA 300  
 AGGAGATATA TTCTATGAAA AGACTTTTAT TTGTGATGAT AGCTTTCGTT TTCATATTGG 360  
 40 CTGCATGCGG AAACAATTTC TCGAAAAGACA AGGAAGCTAG TAAAGATAGC AAGACAATTA 420  
 ATGTTGGGAC TGAGGGGACT TATGCACCAT TTAGTTTCCA CGATAAAGAT GGTAAATTAA 480  
 CTGGTTACGA TATTGATGTT ATTAAGAGCAG TGGCTAAAGA AGAAGGTTTA AAACTTAAAT 540  
 45 TTAATGAAC TTCTTGGGAT TCTATGTTTG CAGGTTTAGA CGCAGGCGCT TTTGATGTAA 600  
 TCGCGAACCA AGTAGGTATT AATCCTGATA GAGAAAAGAA ATATAAATTT TCTAAGCCTT 660  
 50 ACACATTCTC AAGTGCTGTT TTAGTTATTTC GTGAAAATGA AAAAGATATT AAAGATTTTG 720  
 ATGATGTTAA AGGTAAGAAG TTAGCACAAA CATTACATC TAATTATGGT AAATTAGCTA 780  
 AGGATAAAGG TGCTGATATT ACAAAGTTG ATGGCTTTAA CCAATCAATG GATTATTAT 840

55

|    |            |             |             |            |             |             |      |
|----|------------|-------------|-------------|------------|-------------|-------------|------|
|    | AAAAACCTAA | TGCTAAGATC  | AAAGCAATCA  | AAGGTAATGC | TGAACAAAGT  | AGATCTGCAT  | 960  |
|    | TTGCATTITC | TAAAAAAGCA  | GATGATGAAA  | CAGTTCAAAA | ATTCAATGAT  | GGCTTGAAAA  | 1020 |
| 5  | AAATCGAGGA | AAACGGTGAA  | TTAGCTAAAA  | TAGGTAAGAA | ATGGTTTGCT  | CAAGATGTTT  | 1080 |
|    | CTAAATCTAA | ATAGCGAACA  | ACAACATGCG  | CTAGATGCTG | CAAAACAAGC  | TTTCGGACCT  | 1140 |
|    | ATGCTAGAAG | GTTTGGTCAA  | ATATTCAATT  | CCTATTACAT | TAGTTACATT  | TGTTTTAGGA  | 1200 |
| 10 | TTGATTAITG | CATTATTTAC  | AGCAATTAATG | CGAATTTCAA | CGAGTAAAAA  | TTTAAGAAAT  | 1260 |
|    | ATTTACAGTG | TCTATGTATC  | TATTATTGGA  | GGAACACCAA | TGATAGTACA  | ACTATTTATC  | 1320 |
| 15 | ATATTTTATG | GTATTCCAGA  | ATTAGGTAGA  | TTATTAACAA | ATGACGCTGA  | CAACCAATGG  | 1380 |
|    | ACATTGGCAC | CTGTAGTGGC  | TGCTATTATT  | GGTTTATCAT | TAAATGTAGG  | TGCGTATGCT  | 1440 |
|    | TCGGAAATTA | TTGCTGGCGG  | TATTATTTCT  | ATACCGAAAG | GACAAACAGA  | AGCTGCATAT  | 1500 |
| 20 | TCCATCGGTA | TGACGTATGG  | TCAAACGATA  | CAACGTATCA | TTTTACCGCA  | GGCAATTGGA  | 1560 |
|    | GTGTCGATTG | CTGCACTAGG  | TAATACATTT  | TTAAGTTTAA | TCAAAGATAC  | ATCATTATTA  | 1620 |
|    | GGATTTATTT | TAGTGGCTGA  | AATGTTTAGA  | AAAGCTCAAG | AAGTTGCGTC  | TACAACATAT  | 1680 |
| 25 | GAATATTTAA | CAATTATGTG  | GTTAGTTGCG  | CTAATGTACT | GGGTGCTATG  | CTTTATTATT  | 1740 |
|    | TCAATTATCC | AAGGTATCTA  | TGAATCTTAT  | ATTGAAAGAG | GGTATCGCTC  | ATGATTTCAAT | 1800 |
| 30 | TGAACAATAT | CCATAAATCA  | TTTAATGATG  | TTGAAGTCAT | CAAAGGTATT  | GATTTATCTG  | 1860 |
|    | TTGAACAAGG | TGAGGTTGTA  | ACCTTAATCG  | GTGATCTCTG | TTCAAGGTAAA | ACAACATTGT  | 1920 |
|    | TACGTATGAT | TAATGCATTA  | GAAATTCCAA  | CTGAAGGTAC | AGTTTATGTT  | AACGGCAAAA  | 1980 |
| 35 | CATATACATC | TAAAGATAAA  | AAATCACAAA  | TAGAAGTTGC | TAAACAGTCT  | GGTATGGTAT  | 2040 |
|    | TTCAAAGTTA | TAACCTTTTT  | CCGCATAAGA  | CGGCATTAGA | AAATGTAATG  | GAAGTCTTTA  | 2100 |
|    | TCAAGGTATA | AAAGTTGAAA  | AAGGATGAGG  | CACGTGGGAA | ATCACTTGAG  | TTACTTGAGA  | 2160 |
| 40 | AAGTTGGTTT | AACACATGTC  | AAAGATCAAC  | GTCCACATGC | ATTATCAGGT  | GGTCAACAAC  | 2220 |
|    | AACGTGTGTC | TATGTCAAG   | AGCACTAGCA  | ATGAACCTTA | AAATGATGTT  | GTTTGATGAA  | 2280 |
| 45 | CCAACATCTG | CACCTGATCC  | TGAACTTGTG  | AATGATGTTT | TAAAGGTTAT  | TAAAGATTGT  | 2340 |
|    | GCTAATGAAG | GCATGACAAT  | GGTCATTGTG  | ACACATGAAA | TGCGTTTTGC  | TAAAGAAGTA  | 2400 |
|    | TCTAATAACA | TTGTATTTAT  | TCTGGAAGGC  | ATGATCGGAG | AACAAGGGGC  | TCCAGAAGAG  | 2460 |
| 50 | ATGTTCAATC | GTCCGAAAAAC | AGAGAATTA   | AGACGTTTCT | TAAATGTTAT  | AAATGAAGAA  | 2520 |
|    | TAATCAAATA | GAACCAAGTA  | TCAATGTTTT  | GTATGGCGAT | GAACCCATAT  | ACATGATGCG  | 2580 |
|    | TGGTTCTTTG | TTATGTTGTC  | ATAATCTTGG  | AGCGATATTT | TAAACGAGTT  | TATGATTTAA  | 2640 |

TTCTACATGT GCGTTAAAC CTFTTTTGAA TTGTTGGACG CCATAGTCTT CTGATGACTC 2760  
 TGAAGAAGTCA CCGGTAATAC CATAAAAATT ATAGCGATCA ATATGATGCG CTTTAGCAAA 2820  
 5 CTTAATCATT TCCCACTGCA AATGGTAGGC ACCCATATAA GCATTATATT TAGGGTTTGA 2880  
 ACCACTAGAT AAGTAATAAA CTTCATGCTC ATTGTAGATA AATAAAGCAG AAGCTAAGTT 2940  
 10 TAAGACTGCA CCATCTTGTT CAATTGGTTC TATTGTATTG TCGATTTTAC GCTTATTGCT 3000  
 ATTTAGCTGT TGTCTAGCT GTGTGCGTTT CGTTTATT TTCTTTGAAT TAGGACTTTC 3060  
 TTCCAATGCT TCTTCAACAC CTGAGAGTTC AGCTGTTAAT TGTGTGTGCT TTAATTGTAA 3120  
 15 CGTTTTTAAA TACTCGTTTA AATCAATATA CGCCAACTTT AACATGGCGT GGTTCATCGTA 3180  
 TAACTTTTCG ATTCTTCAA AGTATGGTAA CTCACGGAAT TTGAAACCGT GCTTTTCCTC 3240  
 AGCCATATGG AATAAGTCGA AAAAAGTTTG CGTTTCATCA ATCGTTAAAG TTTTAgTTTT 3300  
 20 GACACCAAtA TcATATGtTt tTtTAATATT ACGTCTCGtT TGATAATCCA TTTCTTTTAA 3360  
 AAGTTGGTCT TcAGCTTTAT CTTTTAAATC TAACACTGAC AGCCAACGGA TTTGGCTCAT 3420  
 TGAATCATAA CCTACAGGGA AACCTTGGTG TTTATAACCT AATTATCCA TTGTTCTaAC 3480  
 25 AAATGCTCGG TTATCATAAG ATTAAcAAT TTCACCGTCT GCATTGCGTA AATTTTCAAT 3540  
 TAAATATGGA TCTACAAGGA CATATAAACA ATTGTGTTTC TTTAAATATG ACGTTAATGC 3600  
 30 TTTAAGAGAA AATGCTACTA ATGATTGATT TGTATAATCC ATCACTGGCC CGCGATGTGT 3660  
 ATAAAAATAT TTGAAAAATT TAAGTGTGCG TGCTTCtGTC AATAAGCATC CTGCAATCAC 3720  
 TTGACCATTG TCATCTTTAA CCCCTACrAG ATGCACATCG CCTTTTAAAT CAACTCTATG 3780  
 35 ATTGTAATGA ATAGCTGATT GTGTGTAAATG TGAAAAATGC T 3821

(2) INFORMATION FOR SEQ ID NO: 305:

(1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1422 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 305:

GCGACCATCAT TGTTTGCGAG AACTGTTAAT TTACCAGTTG ATGGTTTCGA CGGTGTTGTT 60  
 50 GGTTTTGATG GTGTAGTAGG GGTAGTTGTA GATTGCGTAC CCCATGGCGC CACTTTACCC 120  
 ATTTTTATTA AATATTTTTC ATTAATTAG TCATATAATT GATCATAACT ATAATTATGA 180  
 CTTCTTAAAT ATCCATGTGG ATCGGCATGG TCAGTACCAC CTAATATATT ACTTACAGCG 240





|    |  |      |
|----|--|------|
|    | GTAGATAATC GAATTATTCC TAAAAATATA ACTCAAACA CAATCTTCAA ATTGAGTAAT   | 360  |
| 5  | TTAACCTTAG AGGTAAACCT CTATGAAAGA AACACAAATT ATCACGGAGA CTTTGCTATT  | 420  |
|    | TACTATCCTG TAGAATCTGC TCTAATAATA GAAACAGATA CTCCTAGGCT AATTAATCAT  | 480  |
|    | ATAAATAACA ACCCAGTCCA AAAAATATTT ATAATAACAA CAATGACTG GAGCTTTAAT   | 540  |
| 10 | ACACTTGaAT TAGAAAAAT AGTAGACGAA ACTATAATTT ATGATTTAAA ACAAGAAGAT   | 600  |
|    | CCTAAACAAT ATGAAATTTT ATATAAAAT AAACACGAA AATTGCCTTA TTGATGAATT    | 660  |
|    | AAACGCAAA CTTAGAACAC CTTACCCCTT TTCCGTCCTT ATTGTAAATT TCCAACTACT   | 720  |
| 15 | CAAAATTAAA ACCACCGTG CAAACGGTG GTTTTATTAT ATATAATTAT TCCCACTCAA    | 780  |
|    | TTGTGCTTGG TGGTTTGGAT GTAATGTCAT AGACTACGCG GTTGACGTGA TCTACTTCGT  | 840  |
|    | TTACGATACG ACTAGAAATC TTTTGTAAGA CTCCCAATC GATGCGTGGC AAGTCACTTG   | 900  |
| 20 | TCATACGGTC GATAGATGTT ACTGCAGGAA TACCTACTGT GTGATCATAC GTACGGTAGT  | 960  |
|    | CTCCATAAC ACCTACTGAT TGAATGTTTG GTAACACTGT GAAGTATTGC CAAATTCTC    | 1020 |
|    | TTTCAAGACC TTCTTCTCTA ATCACTTGGC GTAAAAATCG GTCTGATTCT CTAACGATTT  | 1080 |
| 25 | CTAGTTTATC TTCAGTAATT TCTCCAAGTA CACGAATACC AAGACCAGGT CCTGGGAATG  | 1140 |
|    | GTGTCTCCA TACTAAATGT TCTGAATAC CTAACCTCAAT ACCTAATTTA CGTACTTCAT   | 1200 |
| 30 | CTTTAAACAA TGTATTGATT GGTTCGATTA ATTCGAATTC CATGTCTTCT GGTAAATCCAC | 1260 |
|    | CAACATTGIG GTGTGATTGG AITGTTTGTG CTGTTTTAGT ACCTGATTGG ATGACGCTCG  | 1320 |
|    | TATATAGTGT TCCTTGCSCA AGGAAGTCTA CACCTTTCAG TTTTGATGCT TCATCATCAA  | 1380 |
| 35 | ATACGTATAC AAATTCATTA CCAATGATTT TACGTTTTTG TTCAGGATCT GAAACACCTT  | 1440 |
|    | TTAATTTATT CATAAAGCGA TCTTTTGCAT TAACACGAAT AATATTATG TTGAAACCTT   | 1500 |
|    | CACGGAATTG CTCCATAACC ATGTGCGCTT CACCTTTACG AAGCAAGCCA TGGTCTACAA  | 1560 |
| 40 | AGATACATGT TAGTTGATCA CCTATTGCTT TATGCAATAG TACAGCTACA ACAGATGAAT  | 1620 |
|    | CTACGCCGCC ACTCATGCA CATAATACAC GACGGTCTCC TACGCGTTGA CGAATCTTTT   | 1680 |
|    | CAATTTGAT TTCGATAAAG TTTTCCATTG TCCATTGACC TCTACAATCA CAACACGAC    | 1740 |
| 45 | GGACAAAATT ATTTAATAAA TCATTACCAT ATTCTGTATG ACGTACTTCT GGATGGAATT  | 1800 |
|    | GAACACCATA AATGCGACGT TTCTTATCTT CGATTGCTGC ATAGTCTGTG CTTGGGTAT   | 1860 |
| 50 | CAGCGATAAC TTCAAAGCCT TCTGGAATTT CAATAACTTT ATCAGAATGA CTCATCCAAA  | 1920 |
|    | CAGTTTGTTT TGCTGGTAAG CCAGCGAATA ACTCATCTGA CTTGCGATTA ATGATTGCTT  | 1980 |
|    | TACCGTATTC ACGTTCAITG GCACGTTCAA CTTTACCACC TAATAATTTA GTAGTTAATT  | 2040 |

EP 0 786 519 A2

|    |   |      |
|----|---|------|
|    | TAAATGAACC TTCTTCATAA ACTGAATTGG GACCACCTGA TAAGATAATA CCTTTTGGAT   | 2160 |
|    | TCATTTTCTT AATTTCCTCA ATTGAAATTT CATGATCGTG TAATTCACCTA TAAACGCCCA  | 2220 |
| 5  | TTTCACGAAT TCGGCGTGT AATTAATTGGT TGTATTGGCT ACCAAAGTCT AAGACAAGGA   | 2280 |
|    | TTAACTCTTG TTCTTTTGCC ATTTCATAT TTGTCGTCT CTTTATCTT AATTAGAATG      | 2340 |
| 10 | AGTAGTTCGG TGATTCCTTC GTAATTGGAA TATTATGTGG ATGGCTTTCT GCTAAACCCAG  | 2400 |
|    | CAGGACCCAT ACGTGTAAT TGTGCTTCTT CGCGTAATTC TCTTAAATCG TGTGAACCCAG   | 2460 |
|    | TATAAACCCAT ACCAGCACGC ACACCGCCCA TTAATTGGTA AATTGTATCT TGTAAACGCAC | 2520 |
| 15 | CTTTATAAGC CGTACGCTCT TCGATACCTT CAGGAACAAA TTTCTTAGGC GCTTTGTCTT   | 2580 |
|    | CTTGAAGTA ACGTGTGTTT GAACCTTTT CCATCGCACC TAAAGAGCCC ATACCACGGT     | 2640 |
|    | ATACTTTATA TTGTCTACCT TGGAAAAATTT CTGTGCGCC TGGGCTTTCT TCAGTACCTG   | 2700 |
| 20 | CTAATAAGCT ACCTAACATA ACCGCATGTC CACCAGCAGC TAATGCTTTA ATGATATCTC   | 2760 |
|    | CTGAGAATTT AATACCACCA TCAGCAATGA TAGCTTTACC ATGTTTGGCT GCTTCAGTGG   | 2820 |
|    | CACAATCATA AATTGCTGTA ATTTGTGGTA CACCAACACC TGCTACAACA CGCGTCGTAC   | 2880 |
| 25 | AAATTGAACC TGGGCCAATA CCAACTTTAA CAATATCTGC ACCCGCTTCA AATAAATCTT   | 2940 |
|    | TTGTTGCTTC TGCAGTTGCT ACGTTACCTG CTACTAATGT GATTTCGGG TAAGTCTTCT    | 3000 |
| 30 | TAATATGTTT CACTTGATCG ATAACACCTT TAGAGTGACC ATGTGCTGTA TCGATAACTA   | 3060 |
|    | AGACATCCAC ACCTGCTTCG ACTAATTTTT GAGCAGGAAT ATCAGTATCT TTTGAAATAC   | 3120 |
|    | CAATTGCTGC GGTACAAGT AGACGACCAT GTTCATCTTT TGCTGCATTA GGAATTCGA     | 3180 |
| 35 | TAACTTTTTC AATATCTTTA ATAGTAATAA GACCTTCTAG ACGTCCGTCT TTAATAATG    | 3240 |
|    | GTAACTTTTC AATCTTATGT TTTTGAGAA TTTTCTCTGC TTCTTCAAGT GTTGTATTCA    | 3300 |
|    | CTGGAGCTGT AATTAAATTC TCTTGCCTCA TTACATCTAC AATTTTAAAT GAGAAGTCTT   | 3360 |
| 40 | CAATAAAACG TAAGTCACGG TTTGTTAAAA TACCTACTAA GTTGCATCT TCTTTATTAT    | 3420 |
|    | CAACAATTGG TACACCTGAA ATACGGTATT TACCCATTAA TGCTTCTGCT TCATAAACGC   | 3480 |
|    | TTTCTTCTGG CGTTAAGAAA AATGGGTTTG AAATGACACC ATTTTCTGAG CGTTTTACTT   | 3540 |
| 45 | TTTGAATCTC GTCCGCTTGT TCTTCAACGC CCATATTTTT ATGAATAACA CCTAAACCCAC  | 3600 |
|    | CTTGACGAGC CATAGCAATC GCCATTTTAG ATTCAGTTAC AGTATCCATA CCAGCAGAAA   | 3660 |
| 50 | TAACTGGAAT ATTTAATTTA ACTTTGTCTG ATAATTGTAC GCTTAAATCA ACGTCTTTGG   | 3720 |
|    | GTAATAATATC AGATTGTGCT GGAATTAATA ACACATCATC AAACGTTAAT GATTCTTTTG  | 3780 |
| 55 | CAAAATTTACT TTCCACATT AAAAACAGCC TCCATTTTTT AAATTAATTA GTTATATTAT   | 3840 |

EP 0 786 519 A2

|    |             |             |            |             |            |            |      |
|----|-------------|-------------|------------|-------------|------------|------------|------|
|    | GCAGAGATTG  | CGCCTAAAC   | AATTCGGTTT | TGAGTCAACC  | ATGCAAATTG | TTCACCTAAA | 3960 |
|    | CCTTAATAATG | CTTGTTGGTAC | AGCGCTTATA | CCAGTACCTA  | ATCCTACTGA | TACAGCGATA | 4020 |
| 5  | ATTAATAAAT  | TGTTTTGATT  | TTTAAATCG  | ATATGTCTTA  | ATATACTAAC | ACCATATGCC | 4080 |
|    | ATTACCATGC  | CAACATAGC   | TATCATCGCA | CGCCTAACA   | CAGGTAGCGG | TATGATATTT | 4140 |
| 10 | GCTAATGCGC  | CAAGCTTAGG  | TATACAACCA | CATATAAGTA  | ATAACACGAC | CATGCCGTAT | 4200 |
|    | ATAACATTGT  | TTTTCTTAGC  | GCCGGATAAA | GAACAAGTC   | CTACATTTTG | CGAATAGGCT | 4260 |
|    | GTATACGGAA  | ATGAATTGAA  | TATAGAACCT | AACACTATCG  | CTAGACCTTC | CGCAGTATAA | 4320 |
| 15 | CCTTTACGAA  | AATCTTTTCT  | TTCTAACTTC | TTACCGGTAA  | TTTCACTTAA | CGCATGATAG | 4380 |
|    | ACACCTGTGC  | ACTCAATTAA  | ACTAACGATA | GCTCAATAAA  | AGAACACTAA | CGTCGATGTC | 4440 |
|    | ACATCAAAGC  | TAAATCCAGA  | GAATCTAAAC | GGCACTGGGA  | TGCCTAACCA | ACCGGCATGA | 4500 |
| 20 | TTGACTTGAT  | TAATATCGAC  | CATCCCAAGT | AAGCCAGCAC  | CTATCGTTCC | TAAACGAGT  | 4560 |
|    | CCAATTAATA  | TGGCAATACT  | CTTAATAAAT | CCAGTTGTGA  | ATCTTTGTAA | AAGAAGAATA | 4620 |
|    | ATGATTAATG  | TCATTAAACC  | TAACAAAATG | TTCTTAAACAT | CTCCATAGTC | CTTTGCACCT | 4680 |
| 25 | TGACCTCCAG  | CTAAGTAATT  | CATTGCTACT | GGCATTAAAT  | TGATACCAAT | GATAGTAAAC | 4740 |
|    | ACACTACCCG  | TTACTACTGG  | TGGGAAGAAT | TTTACAAGAT  | GTGAAAAGAA | AGGCGCGATG | 4800 |
| 30 | ATAATAACTA  | ATATCCCTGA  | TAAAAATAGC | GAACCATAAA  | GTACATCTAT | TCCTTTCGTT | 4860 |
|    | TGACCAATTA  | AAATCATGGG  | CGCAACAGCC | GTGAATGTAC  | ATCCAAGAAC | GATTGGTAAT | 4920 |
|    | CCTGTTCTGT  | TTACTTTTAT  | GGCTTGTAAA | AATGTGGCAA  | CCCCACACAT | AAATATATCT | 4980 |
| 35 | ACTGTAACTA  | AGTAAGCGAT  | TTGTTTCAGT | GTAAACTTCA  | AACCTGTACC | AACAATGAT  | 5040 |
|    | GGAACATAAG  | TAGCACCTGC  | GTACATAGCT | AAAAGATGTT  | GAACACTTAG | GATTAAATTT | 5100 |
|    | TTCTATTATC  | TTCTCCACC   | AATGTCACTT | TGTTTCTTC   | TAGTGAAGCA | ACCTTGCAGA | 5160 |
| 40 | GAGAAGAAAC  | TGTTAAACCT  | GCTTCTTCTA | AACGTTGATG  | CCCATTTTGG | AAACTCTTTT | 5220 |
|    | CAACAACAAT  | ACCAATACCA  | GCTGTCTTAG | CATTGCTTGG  | CTGTGCGATA | TCGTATAATC | 5280 |
|    | CTAATGAAGC  | ATCACCATTT  | GCTAAAAAGT | CATCGATGAT  | AAGTACAGTA | TCTTCTTCTG | 5340 |
| 45 | ATAAAACTC   | TTTTGAAACA  | ATGACCGTAC | TTGTTTTATT  | TTTAGTAAAT | GAATGAATAG | 5400 |
|    | ATGTTTCATA  | ATAACCATCC  | GTCAAAGTGC | TAGGTTTTGC  | TTTTTTCGCA | AATAAACATG | 5460 |
| 50 | GCACATCAAA  | ATGCAGTGCA  | GCCATGATTG | CAGGTGCGAT  | ACCGGAAGCT | TCAATGGTTA | 5520 |
|    | AGATTTTATG  | AATCCCTTIA  | TCTTTAAATT | GCTCGTAAAA  | AGTGCGACCA | ACTTCATCA  | 5580 |
|    | TAACTTTTGC  | ATCAATTGTA  | TGATTTAAAA | ATCCATCGAC  | TTTTAAATTC | TTCTCATCAA | 5640 |





55

TGCCCAATCA ATtCTtTCT TAtCCTCAGC AgTTAATCTA CCAATCCTT TTTGATGTGG 4680  
 AAAACGACCA TATGAACTA ATTCCCCAAC AGTTAAGCCA TCTGCTACTT CaGGTGaTTG 4740  
 5 agGTAAAaTg GctATTtTTTt TgCAATCyCy TtCGTAGAtt GTGtATGAAT ATTTTCACcA 4800  
 TCTAAAAATA CTTCGCTTTC TTTAACTGCC AATAAACGTG ACAATGCCTT TAGCAAAAGTA 4860  
 10 GATTTCCTCG AGCCGTTAGG ACCAATGATT GACGTCACTT TGCCATCTGG TATTTCACAA 4920  
 TCTAATTTAT TTATAATCGT GTTATCCCGT TAACCAATTT TAACtGTtTG TCCATGCAAA 4980  
 CGATTcATAA TtTCCTACT TtCAATAAAA TtCTTtCTGT TtATAAAAAA TAATTTCTAT 5040  
 15 TTTTAAATTA TCAATTTTCA AAGACATCCC AATTGATAAT GATTATCATG AACATCATTa 5100  
 TAACATTTTt CAATCTTATT GACTAACATT ACTTTTtTAA TTGGATAGCT CGATTtTGTCa 5160  
 TGtCTtGTAT ATTACTTTTA TAAATAAAAA AACGCCACA GATAAGTCTT CATAGTTCAa 5220  
 20 AAActTGtCC GTGGACTtCT ATTTAAGTAT GTGTGCTCAT ACCAItTtATT TATTCACTG 5280  
 CAAGAAAGCC ATTACCATAG ACATCTCTTA CATCATGAAT TACGAGGAAT GCATCTTTAT 5340  
 CGATTtGTtT AAtTAATCGC TtTGCTtTTG AAActTGtGT TTTAGAAATA ACAACGTATA 5400  
 25 AGACATCTTt TtCTTCAGC GTATAATAGC CATGTCGGTt TAAATGGTt AAACCTCTtC 5460  
 CAATTtGtCT GtCTATTGCT TTGGCAAGTt TGTCGGGATT AGtTGAAATa ATGCTCATAG 5520  
 CTtTTtTAGT GtTTAAACCT TCTATGACAT ATTCCATCAC TTTTGtTtCT ATATAAAGTG 5580  
 30 ATATTACTGT TACTAATACT TTATCAAGTG GAATAACTGT AAGTGAAATt GCAACAACGA 5640  
 TCATATCGAA GAAAAGCAAA GCATATGCCG TGCTTACATC GAGGTATTTC GTTGCAATTC 5700  
 35 TCGCCAAAAA TGtTGtACCT GCTGTGTAC CGCCTGCAAG GATAATTACT CCGATTCTCTa 5760  
 GTCCAACGCT TACACCACCA AAAATGGCAT TCACAATGCT GTTtCCAGTt TCTACTTGCC 5820  
 ATGATTCTGT TAAACTCAAA AATATTGAAA TAAGAATTGT TACAAGAATA GTTAAGTACA 5880  
 40 TACTTCTCTT ACTCAAAAt TTATAAcCTa TGGCAATCAa TACTGCGtTG ACCAAGAAGT 5940  
 TAGTGATGGC TGGTGAAATA TGAACGCAT AATATAAAAT AATTGCTAAA CCTGTAACCC 6000  
 CGCCTTCACC TAAGTTACCA GAAATAAATA ATGCATTAC ACCTGCAGCA AAGATAAATG 6060  
 45 AACCTAAGAC AACTAGTATT AAActTTTAA CCGTTTtATT CAGGAAACCA TCCCCTTTAT 6120  
 ATATTATTA GACTAT 6136

(2) INFORMATION FOR SEQ ID NO: 308:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 2576 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double



EP 0 786 519 A2

| (x1) SEQUENCE DESCRIPTION: SEQ ID NO: 308: |  |      |
|--|--|------|
| 5  | GATATCGTAa CTACTGAAAA GTCATCACCA CCATGGAATG ATTCTTTTAA ACGTTTTAGC  | 60   |
|  | TTCTATGCAA TTGCAATTCA ATACTTTGTT GTACAATTTA TCATTACATT ATTCTTAATT  | 120  |
| 10   | TGGTACCGA CGTATTTAAC AGAAGTATTC CACGTTAACT TTAAGAAAT GAGCATTAGT    | 180  |
|  | TCATTACCTT GGTATTAAAT GTTCTTCTTA ATCTTATCAG CAGGTGCAAT TTCTGACCGT  | 240  |
|  | GTATTAGGAT TAGGTCGTTT AAAATTGCTA GCTAGAGGTG TAAATGCAAT TGCAGGATTT  | 300  |
| 15   | ATTGTGTTTG CAGTTTCAAT TATCTTTGCT GTACGCACAG GAAATTTATA TGTAAGTATT  | 360  |
|  | TTCTGGTTAT CACTAGGTCT TGGTGGTATC GGTATTTCAA TGGGTATGAG TTGGGCTGCA  | 420  |
|  | GCAACTGACT TAGGACGTAA CTTCTCTGGT ACAGTATCAG GGTGGATGAA CTTATGGGGT  | 480  |
| 20   | AATATAGGTG CATTAAATCAG TCCGCTATTA GCAGGTCTAT TCGTAGAACA TTTGGGTTGG | 540  |
|  | ACAATGACAT TCCAATTGTT AATCGTTCCA GCAGTAATCG CTGTGATTAT GTGGTTCTAT  | 600  |
|  | GTGAAACCGA ATCAACCTTT AATTGTTAGT GATGATAAAG CAATAGAAAA ATAATTTTAA  | 660  |
| 25   | CAAGCAGTAA GCTTTCACAT AGTTGGGGCT TATTGCTTTT TTGCGTTGA AATTGAAACT   | 720  |
|  | TTTTAAAAACA GATATGGTTT AAGATGAAAA TGAAGTTATT GAAATGATAT ATGTAAAGAA | 780  |
|  | ATAAGGTTTT AAAACATTAG TCAGGTAACG CTTGTAAAG TACATATAAA TTTTAACTAG   | 840  |
| 30   | CGCAAGGTG GCGACCAAAA GTCaACGAT GTTAAATAAC aTTAGTAATT AATTITTAATT   | 900  |
|  | GGACTTTAAA AGTTTTTAAA TTTAGATAAT TGAGCATAAG GTGTTATAAT GACATATGTT  | 960  |
| 35   | CGGTAATTAA AATTTATAGC AACAAATTCa TTTTAACTAT GCTAATAAAA AGATTATGGA  | 1020 |
|  | AATATTTTGA CAAGGAAAGG AGAAGTCGAA ATGACATCTT TTTGACATCA CTCATAAAAA  | 1080 |
|  | TCAaTCGACT TAACCTTAGC TTTTATAAAG GTGTAAGACA GGGACTGTTA ATGATTATTTC | 1140 |
| 40   | CTGCAATAAT CGGTTACTTA TGTGGTAATT TCCAATTGG ATTATTAGTT GCAACCGGAA   | 1200 |
|  | CACTAGCCCA TATTTATGTT TTTAAAGGTC CGTCGCGATC TAAGCTGCGA ACTGTAATAA  | 1260 |
|  | TTTGTAATTT AGCGTTTGCA ATATGTATGA TGCTTGGTAC GCTAACAGCC AAAACGCCAC  | 1320 |
| 45   | TCGTTTTTGG AATGACATTA TTAATTGTTA CGGTTATACC ATTTTATATA TTTACTGCCT  | 1380 |
|  | TAAAAATAGC TGGACCGTCA TCGACATTCT TCATGTGAC ATTCACTCTA CCCATTAACT   | 1440 |
|  | TACCTATAGC TCCGGAAGAA GCATTATATA GAGGCTTTGC GATTTTAGTA GCGCGTATAC  | 1500 |
| 50   | TGCCCACTAT GATGGTGTTA ATCAGCATCG TATTTTCTAA AAACAAAGCT GAAGAACAAAG | 1560 |
|  | CAATTCAAAA TGATTTTAAA CTCATATCTA AGTTGTTACA CACTTATAAT GATAAATCTG  | 1620 |

EP 0 786 519 A2

TCACTTCTAC TTCAAGTAAC GATAAATTAA GTAGACGTTT CCAAAAATTA TTATTATTAC 1740  
 ACACATCTGC CCAAGGGATT TATTCTGAAC TGTTAGAGTT GAACGCTAAA CAAATTGAC 1800  
 5 CATTGCCAGA TGAGTTAATT GAAATGATGG ATCATATCAT TGCACAACATA GATAATAGTG 1860  
 AGGAAATGT AAGATATTGG CGAAAAGAAG TGACAGTAAC AGAGGAATTT CAAAATTTAT 1920  
 TCAACCATAT ATTGAAAATT GATGAAATGG TGATGCAAAA TGAAGCGCGT ATTGCGTATG 1980  
 10 AAGCAGACAT GCGAAACCT TTATATAGTA AACGCATTTA TCaaaaattta AcaTTAGACT 2040  
 CtAtKgtTTT TAGAAATACA TTGAGATATA CAGCGATTAT GATGATAGCG ATATTTATTG 2100  
 15 GGTAAATGTT TGATTTTGAA AAAGCATACT GGATACCGTT ATCTGCACAT ACAATTATTAC 2160  
 TAGGAACATC AACTATACAT GCAATCGAGA GAGGTATGGC ACGAGGTTTA GGTACTATTT 2220  
 TAGGTGTGTT AGTACTTTCA GTCATATTGT TGTTTTCAAT ACCAACACCT GTTGCGATTA 2280  
 20 TTTTAATGGG CATTGCGACA TTGTTTACTG AAGCATTGGT GCGAGCAAAAT TATGCGATTG 2340  
 CAGTAGTTTT TATTACAATA CAAGTTATTT TAATGAACGG ATTAGCATCA CAGAAATTAA 2400  
 CAATTAAACAT TGC GTTTCCA AGAGTTATTG ACGTTGCAAT GGTATTGTG ATTGCAATCA 2460  
 25 TAGGTTTATT TGTCCTTGGA CAACGTACCG CATCCGCATT GCTTCCTAAT GTAATTGGCTG 2520  
 AAGTTGTTGG TAAAGAAACA ACGCTCTTTC ATTATTATT TTCTGAAAAT CAATAT 2576

(2) INFORMATION FOR SEQ ID NO: 309:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 668 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 309:

CAAGCACATT ACGATGCTCT AATACTTTGC TAATAATTTT TTCTTGCTCT GGTGAAACG 60  
 TTTCATACCC AAAGTAATGC GATAATGTTT GTTGATCAT AAATTGACCC CTTATTGTTG 120  
 TTCCTTAATT TCTCTAACT CACTCCATCT TGTGATGCT AAATCATATT GAATTTCAAG 180  
 45 TGTTCTTTT TCTTCGTTA ATCTTTAAT TTCCCATAA TCTGCACITG CCTCAATCAT 240  
 GAGCACATCA ATTTCTTCCA TTCTTACTTC CGCTTGTTCT ATGCGTTTCA TCAATTGTTT 300  
 ATATTCTAAT TTTTCTTAT ATGATAAACC ATTTTCTTIA CGTACAGTTG TAGAAGATTT 360  
 50 AGATTGTTGC TTCAATGTGG ATTTATTTTT ATCTAATGAT TTTTATAAC TTTCAATATC 420  
 TTCAAAAGTT CCGATAATCT TTTCCATCTG ACCATCATGA ATAAACCAAT ATGACGTGTC 480

AATATAATCT TCAAGTATTG TTAAAGTCTC AGTATCTAAA TCATTGTGCG GTTCATCTAA 600  
 CAACAGAACAA TTGCGCTGGT GTACGAGTAG ACGTAATAAA TACAAACGCT TTGCTCTCC 660  
 5 ACCAGATA 668

(2) INFORMATION FOR SEQ ID NO: 310:

10 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 12173 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 310:  
 CATCAGTTTA TTGTGAAAGG CAATGCGATC ATTTCATGTA TTTATTGTGT TTGAAACATC 60  
 20 GTTAGATAAC AATAGTGATA TTGCACATTT TAAGCTGAAG ATGGAAGAAA ATTGCGATAAT 120  
 CAAATAAACAA ATGAATTTTA GAAGGTACAA TGACGTTTCA TAATTTAAAT ATAGCTGAAT 180  
 GTGTGCTGTA GTGATGTTCA CTATAGATTT ATATTAAAT ACAAAGACAA AGGTTGTTAA 240  
 25 TTTTATTAA GCGTTAGGTT GAATGTATGA GAATTTTAGA TTTATAATAG AAGATAGAAA 300  
 CGAAAAATTT TCTTAAAGC AGTAATGTTG ACTCAAAATA AGCTATAATA ATGACACTTA 360  
 TTTAATTGAT TAACATTGTC TAATAAATAT CAATATAGAA TATAACTTTC CAATAATGAC 420  
 30 TGAGAAAATC GAAATGTCAG TCTCGAATCA TATAATTAGA AAATTGATTA TTTTCTGICA 480  
 ATTTAGGGTT GAACTATACA TATGATATTG TTAGAATATT TTTTAACATT ATATTTTATT 540  
 35 GCTTTAAAGT GGAATATACT TGAAATAATT AGTAGAGGTG AGTAAGGATG AGTAATAAAT 600  
 TAGAATCATA CAGAAGTGAG ATTGTATCAC TGAATCATCA AATTTTAGAC TTATTATCTA 660  
 AACGTGGTGA ACTAGCACAA AAAATTGGGG AAGAAAAATT AAAACAAGT ACACGTATTT 720  
 40 ATGATCCACA ACGTGAAAAA GAAATGCTTA ACGACTTAAT CGATAGTAAC AAAGGACCAT 780  
 TCAACGATAA TACTATTAAG CAATTATTTA AAGAAATTTT CAAAGCCTCT ACAGATTTAC 840  
 AAAAATCTGA AAATGAAAAA CATTATATG TATCAGTAA GTTGAAACCT GAAGATACGA 900  
 45 TTGTAAACATT TGATAATGGG GGCATTATAG GAGACGGCAA TAAATCATT GTATTTGGGC 960  
 CATGTTCACT TGAATCATT GAACAAGTTG AAGCTGTTGC TAAAACTTA CATGCTAAAG 1020  
 GTGAAAAAT TATTGTTGGC GGTGCATTTA AACCACTAC ATCACCATTAT GATTTCCAAG 1080  
 50 GCCTAGGTGT TGAAGGACTT AAAATACTTA AACAGATTAA AGATAAATAT GATTTAAATG 1140  
 TTGTCAGCGA AATCGTAAAT CCAAATGATT TTGAAGTGGC TGATGAGTAT TTAGACGTAT 1200

55

EP 0 786 519 A2

|    |   |      |
|----|---|------|
|    | AAAAGCCTAT TCTATTAAAA CGTGGTTTAT CTGCTACAAT CGAAGAGTTT GTTTATGCAG | 1320 |
|    | CTGAATACAT TGCCTCACAA GGTAAATCAA ACATTATTTT ATGTGAACGT GGAATCCGAA | 1380 |
| 5  | CTTATGAAAA GCGACACGT AACACTTTAG ATATTTTCAGC AGTACCAATT TAAAAACAAG | 1440 |
|    | GTACACACTT ACCAGTCATG GTAGATGTTA CGCATAGTAC AGGTCGTAAA GATATCATGT | 1500 |
|    | TACCAACTGC GAATsAgCAT TAGCAGTTGG TGCTGATGGA GTTATGGCTG AGGTGCATCC | 1560 |
| 10 | AGATCCATCT GTTGCACTTA GTGATGCGGG TCAACAAATG GATTIAGATG AATTCCAAGC | 1620 |
|    | ATTTTATGAT GAATTAAAGC CTTTAGCTGA TTTATATAAC GCTAAAAAGT TAAAAATAA  | 1680 |
|    | TTCCAAGGAA ACTATAGACT ACTTAACATA TATGTCATGT TGAAGTAGAA TATTATCTTT | 1740 |
| 15 | GAATCGACAA TTTTAACTT ACAGCCATTC TAAGAGTATA TTACTTTTAG AGTGGCTATT  | 1800 |
|    | ATTTTTTGTA TAGAAATAAA GGTATACCTG ACTTAACGAT TGTATAATA CTTGCACT    | 1860 |
| 20 | TGTTCAATTT CACAATTATT AAAGATTATG ACTGATAGCA GTAATTAAAA TTATAACTAT | 1920 |
|    | GAATTATCTG TAAATATAAA TAGATTACA CATTGTGTGC TGAATGTGA ACATTTTTC    | 1980 |
|    | ACAATGCAAA TTGATATTTG AAAAGGCTTT CTCAAAACAT TACAATTAAA AATGAAAAAA | 2040 |
| 25 | GTTTATATAA AATTAAAAATA TATGTTCTGT TATCATTTAG CGTTTGTGTT TATTTCAGC | 2100 |
|    | TTTTGCTAA ATTTTCCAA ACAAAAATAT GTTACTGTAA ATTAAAAATAT GGTAAACTAT  | 2160 |
| 30 | GAAATGAAA TGAAAACATG TTATTATAAT GAATAAAACG TTTACAAGGA GGAAATTATG  | 2220 |
|    | ACAGTTACTA TATATGATGT AGCAAGAGAA GCGGTGTCT CTATGCCAC AGTGTCGCT    | 2280 |
|    | GTTGTAAATG GGAACCAAAA TGTTAAAGCA GAAACTAAAA ATAAAGTTAA CGAAGTCATT | 2340 |
| 35 | AAGCGTTTGA ATTATCGTCC AAATGCTGTT GCTAGAGTT TAGCTAGTAA AAAGACAACA  | 2400 |
|    | ACAGTAGGTG TGATCATTC AGATATATCT AATATCTATT ATTCACAAC TGTCTGTGGA   | 2460 |
|    | CTTGAGATA TTGCAACAAT GTATAAATAT CACTCAATTA TTTCAAATC AGATAACGAT   | 2520 |
| 40 | CCTGAAAAGG AAAAAGAAAT TTTTAATAAC TTATTAAGTA AACAGGTTGA TGGTATTATT | 2580 |
|    | TTCTTGTGTT GTACAATTAC TGAAGAAATG AAAGAATTGA TAAATCAATC ATCTGTACCT | 2640 |
|    | GTAGTAGTAT CAGGAACAAA TGGTAAGGAT GCACATATAG CATCAGTTAA TATTGATTTT | 2700 |
| 45 | ACTGAAGCTG CGAAGAAAT TACGGGAGAA TTAATTGAAA AAGGCGCTAA ATCATTTGCT  | 2760 |
|    | TTAGTAGGTG GAGAACATTC TAAAAAGCT CAAGAAGATG TTTTAGAAGG TTTAACTGAA  | 2820 |
|    | GTGTTAAATA AAAATGGCCT TCAATTAGGT GATACATTGA ATTGTTCTG TGCTGAAAGT  | 2880 |
| 50 | TATAAAGAG GGTAAAGC TTTTGCCAAA ATGAAAGGCA ATTTGCCAGA TGCCATTTTA    | 2940 |
|    | TGTATCAGCG ACGAAGAAGC AATTGGTATT ATGCATAGTG CAATGGATGC TGGTATTAAA | 3000 |

55

|    |   |      |
|----|---|------|
|    | CCACAACCTT CTAGTGTIAT TCAACCATTA TATGATATCG GTGCAGTAGG GATGCGCTTA   | 3120 |
|    | TTAACAAAAT ATATGAACGA TGAAGAGATA GAAGAACCAA ATGTAGTTTT ACCTCACAGA   | 3180 |
| 5  | ATTGAATACC GAGGAAC TAC AAAATAAATT CACAAAATTA GGCATTCA TC TAACGACCCA | 3240 |
|    | AATTATATGG GTGTTGGAAG AATGCCCTTT ATTTATCTTT TAAAATCGTT GCAGATTAGG   | 3300 |
| 10 | TTACTTATGG ACGAGTAGAT TCGTACCAAC TCGTATATG TAAAGCTAAT TTTTATTTTT    | 3360 |
|    | TTTCACTAAT TTCTTTTGTG CGGGGACAT AGGTATAATC ATTTAAACGA TCTTCCCATC    | 3420 |
|    | TTTTAGGTAA TAATTCAGAT GAATAATGTT TCCATTATTT AACCGTAAAT              | 3480 |
| 15 | AACCACCTTG AATTGGTTGA TCAATTAAC TTAAGAATAC ATGACTCCAT GCACGTGGTA    | 3540 |
|    | CGACTCTCCA AATATTGTAG CCTCCGCCAC CAAACATAAT TACCTTTCCA TTCGTATAAG   | 3600 |
|    | AATCAGCTAA ATATTTTACA AAATATGGAA TTTCAATAAA TGAATGTAAC GTACAATTTA   | 3660 |
| 20 | GATGAGTTAG TGGATCACGA TAATGTATAT CGACACCATT TACGCTTAGA ATAATATCAG   | 3720 |
|    | GTTTAAACT CTTTACGACA GGCTCAACTG TTAATTTAAA ACATCCAAA AATGATGCAT     | 3780 |
|    | CTTCTGTATA CGGTTCAGT GGGACATTTA CAGTGTGTCC ATAGCCGATA TCTTACC GC    | 3840 |
| 25 | GCTCAGTATA GTGACCAGAG CCTGGGAAAA GAAATTTTCC GGTTCATG ATAGAATAAG     | 3900 |
|    | TAGTAACATG GTTATCGGCA TAGAACTCC ATTGTGTACC ATCTCCATGA TGTGCATCG     | 3960 |
| 30 | TATCTATGAT TAAAACGCGT TGATTGTATT CTTTAGCTAA GTATTGTGCG GTAATTGCAA   | 4020 |
|    | TATCATTTGA TATACAAAA CCACTTGCTC GACCAAGTTG AGCGTGATGC AAACACCCAC    | 4080 |
|    | CTAAGTGACA ACCATTTAAT ACTTTCCTG ACATAATAAG ATCTGCTAAA GTTAAAGCGC    | 4140 |
| 35 | CTCCAACAAT TGTGGCACTA TGGCGGTGCA TATGCTTAAA TTGACCATT CTTTCATCAT    | 4200 |
|    | TTAATCCATA TTCTTAGGCC TCATCTTCAC TGATAATGCC ATGTGAAGCA TGCTTAATAG   | 4260 |
|    | CTTCGACGTA ATCATATTTA TGAATTAACA TTAATTCGTC ATCTGTTGCA ATTCTAGGTT   | 4320 |
| 40 | GTACTATTTG TTCTGGAGAC AATAAATTTG CATTCAAAAG TAGCTCTGTT GTTAATTTTA   | 4380 |
|    | AACGCATTGG ATTGAAGGGA TGTGGTGCAT GAAATCGATA TTGTAATAAC TTATCTGAAT   | 4440 |
|    | AAACATATGC AGTTTTTGAT GAATGTTGTT GCATATAATC CCTCCGATAT TCCAAAAATT   | 4500 |
| 45 | AAAAGAAAA CCGATTCATA TAACGAATAT CATCAAAACG TTGTGCTGT TCTAATGTAA     | 4560 |
|    | TGTTTTTGCC AATCTTGCC ATTAACAAT TAGCTGATG ACTTGTTATT TCTGGATCAT      | 4620 |
| 50 | CTGTAGCGAA TATTCAAGT CCACCAAGTT CCATTAAACG CTGCATTAA TTTTATATGT     | 4680 |
|    | CAAATACATC TAACTTGAA TTTTTTAAAT CCCAATGCCA GTAATATTCT GTAGTTATAA    | 4740 |
|    | CGATATAAAT CTCGAATCTT GGTGTAGAAA GGCTAAGTTG TATCAGCTTT TCTGCAAGTT   | 4800 |

EP 0 786 519 A2

TACCTGTAGA CCAGCGCTCA ATTTTCATCAG GATAGTGGAA AGTGACATAA CCCACAATGA 4920  
 GTTGATCTTG GOGAGCAACA TAAATTCCTAC CTTCGTGTAA TGTGTAAAT TCTAACAAAG 4980  
 5 CTTTATACTG ATCTTCAGCA TCTCTAAATG CGGTAAATG CGCATCGAAA GTAAGCGCTT 5040  
 TCAATCTCTC GTGTGTTAAA GGACCTTCAA TAACAAATTG CTGTGCATGA ATGTAATAAT 5100  
 10 CTTCGGATTG ATACGCTCTT AAATGATTCA TATTTTCAAC TCCTCAATCG ACGTGGTGAT 5160  
 TGTATTAAAT TCATTATATA GAAaATTTC AATAATTAAT ACTAGAAAAA GGATAAAAGT 5220  
 AAAAATTTTG AATAATTAGA AATGTTATGT ATAATATTGA GAAAGAAAGC GTTTTCACAT 5280  
 15 AACAAAGGGG GAGTTTCAAA TGAAAGTCGA AGTTTATAAA GGAGCGCAAG GTAAACATAA 5340  
 CCTTAAAGAT TATGAAGAAA CATATAATAC TTTTGATTGG AAAGACGTAG AACAAAGCATT 5400  
 TTCTGGAGT GAAACTGGAA AAATGAACAT GGCATATGAA TGCATAGATC GCCATGTAGA 5460  
 20 TCAAGGATTA GGGGATAAAA TAGCGTTAAA TTACAAAGAT GAGCACAGAA AAGAATCGTA 5520  
 TACTTATAAA GATATGCAAC GGTATCTTAA TAAAGCAGCG AATGTTTTGT CTGAACATGC 5580  
 AGAAGTTGAC AAAGGTGACA GAGTATTTAT ATTTATGTCG CGTACACCTG AACTATATTT 5640  
 25 TGGTTGTTA GGTGTTTTAA AAATTGGTGC AATTGTTGGG CCGTTATTTG AAGCATTTAT 5700  
 GGAAAGGGCA GTTGGCGATA GATTAGAGAA CAGTGAAGCT AAAGTGTTAA TTAATAATAA 5760  
 30 GGCATTGTTA CCTCGAGTAC CTGTAGATAA ATTACCAAACT TGAaaaaaaa TTGTGTCGT 5820  
 AGATGAGGAT GTAGAAGACA ATTACATAGA CTTCAATTAG TTGATGGAAA CTGCTAGCGA 5880  
 TGAATTTGAC ATTGAATGGT TAAAGTCGGA TGATGGTTTG ATTTTACATT ATACATCAGG 5940  
 35 TTCTACTGGG CAACCTAAAG GtGTATTGCA GTTTCACAA GCAATTGTAG TGCACATAT 6000  
 TTCTGGAAAA TATGTATTAG ATTTACAAGA AGATGATGTT TATTGGTGTA CAGCAGATCC 6060  
 AGGTGGGTT aCAGGAACAT CTTATGGTAT TTTTGCACCA TGGTAAATG GCGCTACAAA 6120  
 40 TTGTATAGCT GGTGTCGCT TTTCGCCAGA ACAGTGTAT AGTATGATTG AAGATTTTAA 6180  
 AGTGACGATT TGTATACGG CACCAACAGC TTTAAGAAAT TTAATGAGTG CTGTGACGA 6240  
 TATTGTTGAG AAATATGACT TGTCACTGTT ACGTTCGATT CTATCAGTAG GTGAGCCTTT 6300  
 45 AAATCCTGAA GTTATAAAAT GGGCGAAAAA AGTATACGGT TTAACGGTGT TAGATACTTG 6360  
 GTGGATGACA GAAACAGGTG GACATATGAT TGTTAACTAT CCAACGATGG ACGTCAAGCT 6420  
 TGGCTCAATG GGCAAACCAT TACCTGGTAT TCAAGCTGCA ATTATCGATG ATGCAAGGAA 6480  
 50 TGAATTACCA CCAATCGAA TGGGCAACCT TGCTATAAAA AAAGGCTGGC CATCAATGAT 6540  
 GTATCGTATC TGGAAGAATC CAGAAAAATA TAAATCATAT TTTATTGGAG ACTGGTATGT 6600

55

EP 0 786 519 A2

|    |             |            |             |            |            |             |      |
|----|-------------|------------|-------------|------------|------------|-------------|------|
|    | TGATGTAATT  | ATGACAGCTG | GTGAACGAGT  | TGGACCATTT | GAGGTTGAGT | CTAAATTGGT  | 6720 |
|    | TGAACACGAA  | GCAGTTGCCG | AAGCAGGAAT  | TATTGGTAAA | CCTGATCCGG | TTCCGCGTGA  | 6780 |
| 5  | AATAATTAAAG | GCCTTTGTTG | CACTGAGAAA  | AGGATATGAA | CCAACAGACG | AATTAAAGA   | 6840 |
|    | AGAAATTCGT  | ATATTGTGA  | AAGAAGGTTT  | GTCCGCACAT | GCAGCACCAC | GTGAAATCGA  | 6900 |
| 10 | ATTTAAAGAT  | AAATTACCTA | AAACACGGTC  | AGGTAAAAAT | ATGAGACGTG | TATTAAGAGC  | 6960 |
|    | TTGGGAATTA  | AATTTAGATG | CTGGGGATTT  | AAGTACAATG | GAATAATGAC | ATGAATGTTA  | 7020 |
|    | TTGAAGATTT  | TTTTCGAAGA | ATAAAGGGTG  | ACAACATATT | TCATGTCAAT | GTTTAAATAA  | 7080 |
| 15 | TCGTTTACTT  | TACGATAAGC | AATATAAAGA  | ACTGTTAACT | TGTGTCAAT  | CATTTCGTAG  | 7140 |
|    | AAAGCATTTG  | AAAATGATGA | CATAACAATA  | ATGGCATATC | TTTATATTGC | TTTTTATTTT  | 7200 |
|    | TAATATGATC  | TTTGGAAGAT | GATTATTTTA  | AATAATAGAA | AAATATAGTT | ATCAATAGTA  | 7260 |
| 20 | TCAAGCGCTA  | AAAGTTGTAT | AATACAAAAC  | TTTAATAAGT | GAATTTATTG | CAAAAATGAA  | 7320 |
|    | AGCGCTAACC  | CGATTTAGTC | GACAAGTTTT  | TAACAGTTCC | TTATTATATG | AATGTAAAGTA | 7380 |
|    | AAAATTTCTT  | AGCTACAAC  | TACATATTAT  | AAATGCATAA | ATTAAACAAA | AAGGGGCGAA  | 7440 |
| 25 | AAAAGTTGAC  | TCATTTATCA | GATTTAGATA  | TTGCGAATCA | ATCAACACTA | CAACCAATTA  | 7500 |
|    | AGGATATTGC  | TGCATCAGTA | GGTATTTTCA  | AGGATGCATT | AGAACCTTAT | GGTCATTACA  | 7560 |
| 30 | AAGCTAAAAA  | CGACATTAA  | AAAATTACGC  | CAAGAGAAAA | CAAAGGGAAA | GTGTTTTTAG  | 7620 |
|    | TAATCGCAT   | GAGCCCAACA | CCAGCTGGTG  | AAGGTAATC  | AACGGTTACA | GTGCGTTTAG  | 7680 |
|    | CTGATGCATT  | CCATGAGTTA | AATAAAAACG  | TTATGGTTGC | ATTAAGAGAG | CCTGCTTTAG  | 7740 |
| 35 | GACCAACATT  | TGGTATCAAA | GGTGGTGCGA  | CTGGTGGTGG | TTATGCGCAA | GTCTTACCTA  | 7800 |
|    | TGGAAGATAT  | CAACTTACAT | TTCAACGGAG  | ATTTCATG   | GATTACAAC  | GCAAAATATG  | 7860 |
|    | CATTGTCTGC  | GTTTATCGAT | AATCATATT   | ACCAAGGTAA | CGAATTAGGA | ATCGATCAAA  | 7920 |
| 40 | GACGTATTGA  | GTGGAACCGT | GTATTAGATA  | TGAATGATCG | TGCACTTAGA | CATGTAAACG  | 7980 |
|    | TTGGGTTAGG  | TGGACCTACA | AATGGTGTAC  | CACGTGAAGA | TGGCTTTAAT | ATTACAGTAG  | 8040 |
|    | CGTCTGAAAT  | TATGGCGATT | TTATGTTTTAA | GTAGAAGTAT | TAAAGACTTA | AAAGATAAAA  | 8100 |
| 45 | TTAGTCGTAT  | TACTATTGGT | TACACTAGAG  | ATCGCAAGCC | AGTTACAGTT | GCAGATTTAA  | 8160 |
|    | AAGTGAAGG   | TGCACTTGCA | ATGATTTTAA  | AAGATGCAAT | AAACCAAAAC | TTAGTACAAT  | 8220 |
|    | CAATTGAAGG  | GACACCTGCA | TTAGTTTCATG | GTGGACCAT  | TGCGAATATC | GCACACGGTT  | 8280 |
| 50 | GTAATCTCAAT | TTTAGCAACT | GAACACGCAC  | GTGATTTAGC | TGATATCGTT | GTAACGGGAG  | 8340 |
|    | CTGGATTITG  | TTCAGACTTA | GGCGCTGAAA  | AATTCATGGA | CATTAAAGCG | CGTGAAGCAG  | 8400 |
| 55 |             |            |             |            |            |             |      |

EP 0 786 519 A2

|    |  |       |
|----|--|-------|
|    | GTGTAGCGAA AGATAATTTA AAAGAAGAAA ATGTAGAAGC AGTAAAAGCA GGAATTGTTA  | 8520  |
|    | ATTTAGAGCG TCATGTTAAT AATATTAAAA AATTCGGTGT AGAACCGGTT GTTGCAATTA  | 8580  |
| 5  | ATGCAATTTAT ACATGATACC GATGCAGAAG TAGAATATGT AAAATCTTGG GCTAAGAGAA | 8640  |
|    | ATACGTACG AATTGCCTTA ACTGAAGTTT GGGAAAAAGG TGGTAAAGGT GCGCTTGACT   | 8700  |
| 10 | TAGCAATGA AGTATTAGAA GTCATTGATC AACCTAATTC ATTTAAACCT TTATATGAAT   | 8760  |
|    | TAGAATTACC ATTAGAGCAA AAGATTGAAA AGATTGTGAC TGAAATCTAT GCGCGTTCOA  | 8820  |
|    | AAGTAACGTT TAGCAGTAAA GCGCAAAAAC AATTAAAAACA ATTTAAAGAA AATGGTTGGG | 8880  |
| 15 | ATAATTACCC AGTATGTATG GCGAAAACAC AATATTCTAT CTCAGATGAT CAAACGTTGT  | 8940  |
|    | TAGGTGCACC ATCAGGATTT GAAATTACAA TTCGTGAATT AGAAGCGAAA ACAGGTGCAG  | 9000  |
|    | GATTTATCGT AGCGTTGACA GGTGCAATCA TGACTATGCC TGGTTTACCT AAAAAACCAG  | 9060  |
| 20 | CAGCATTAAA CATGGATGTT ACTGATGATG GTCATGCAAT TGGGTTATTC TAATAAATCA  | 9120  |
|    | TGTCAATGTT TTAATAAAGA TAAGTAAATA GTTTAATAGA CCGACTGTT GGAGATGCAT   | 9180  |
|    | TATTTCAAGC GTTCGGTTTT TTGCTGTGCT AAAAATAGAT TCAATTTGCG GAATCTAACG  | 9240  |
| 25 | ACAATGTTTG AAGGTGGTTA ATTAATGTAT ATGAAGATAA AAAGTGGGCT TGAAGAATAG  | 9300  |
|    | GAAAGCGATG CAATGAATAT TCCATATTAA AAAAAATTAA TAAATAGGT TGCAATATTT   | 9360  |
| 30 | AATTOGGATG CGCTACAATT AACACTAATA ATTGATATTG ATAATTATTA TCAATTAAAT  | 9420  |
|    | ATAATCTTAT AGGAGTTGTT AACCAATGTA ACAAACATCA CCAAAAATTA AGGCTTTTCT  | 9480  |
|    | ATTCTATTAG AAAATCAACT CTAGGCGTTG CATCGGTCAT TGTCAGTACA CTATTTTAA   | 9540  |
| 35 | TTACTTCTCA ACATCAAGCA CAAGCAGCAG AAAATACAAA TACTTCAGAT AAAATCTCGG  | 9600  |
|    | AAAATCAAAA TAATAATGCA ACTACAACCT AGCCACCTAA GGATACAAAT CAAACACAA   | 9660  |
|    | CTGCTACGCA ACCAGCAAAC ACTGCGAAAA ACTATCCTGC AGCGGATGAA TCACTTAAAG  | 9720  |
| 40 | ATGCAATTAA AGATCCTGCA TTAGAAAAATA AAGAACATGA TATAGTCCA AGAGAACAA   | 9780  |
|    | TCAATTTCCA GTTATTAGAT AAAACAATG AAACGCAGTA CTATCACTTT TTCAGCATCA   | 9840  |
|    | AAGATCCAGC AGATGTGTAT TACACTAAAA AGAAAGCAGA AGTTGAATTA GACATCAATA  | 9900  |
| 45 | CTGCTTCAAC ATGGAAGAAG TTTGAAGTCT ATGAAAACAA TCAAAAATTG CCGATGAGAC  | 9960  |
|    | TTGTATCATA TAGTCTGTGA CCAAGAGACC ATGCCTATAT TCGATTCCCA GTTTCAGATG  | 10020 |
| 50 | GCACACAAGA ATTGAAAATT GTTTCTTCGA CTCAAATGTA TGATGGAGAA GAAACAAAT   | 10080 |
|    | ATGATTATAC TAAATTAGTA TTTGCTAAAC CTATTTATAA CGATCCTTCA CTTGTAAAT   | 10140 |
|    | CAGATACAAA TGATGCAGTA GTAACGAATG ATCAATCAAG TTCAGTCGCA AGTAATCAAA  | 10200 |

55



# EP 0 786 519 A2

|    |  |       |
|----|--|-------|
|    | AGGCAACGAC CAATATGAGT CAACCTGCAC AACC AAAAATC GTC AACGAAT GCAGATCAAG | 10320 |
|    | CGTCAAGCCA ACCAGCTCAT GAAACAAATT CTAATGGTAA TACTAACGAT AAAACGAATG    | 10380 |
| 5  | AGTCAAGTAA TCAGTCGGAT GTTAATCAAC AGTATCCACC AGCAGATGAA TCACATACAAG   | 10440 |
|    | ATGCAATTAA AAACCCGGCT ATCATCGATA AaGAACATAC AGCTGATAAT TGGCGACCAA    | 10500 |
| 10 | TTGATTTTCA AATGAAAAAT GATAAAGGTG AAAGACAGTT CTATCATTAT GCTAGTACTG    | 10560 |
|    | TTGAACCAGC AACTGTCATT TTTACAAAA CAGGACCAAT AATTGAATTA GGTTTAAAGA     | 10620 |
|    | CAGCTTCAAC ATGGAAGAAA TTGAAAGTTT ATGAAGGTGA CAAAAAGTTA CCAGTCSAAT    | 10680 |
| 15 | TAGTATCATA TGATTCTGAT AAAGATTATG CCTATATTCTG TTTCCAGTA TCTAATGGTA    | 10740 |
|    | CGAGAGAAGT TAAATTTGTG TCATCTATTG AATATGGTGA GAACATCCAT GAAGACTATG    | 10800 |
|    | ATTATACGCT AATGCTCTTT GCACAGCCTA TTACTAATAA CCCAGACGAC TATGTGGATG    | 10860 |
| 20 | AAGAAACATA CAATTTACAA AAATTATTAG CTCGGTATCA CAAAGCTAAA ACGTTAGAAA    | 10920 |
|    | GACAACTTTA TGAATTAGAA AAATTACAAG AGAAATGCCC AGAAAAATAT AAGCGGAAT     | 10980 |
|    | ATAAAAAGAA ATTAGATCAA ACTAGAGTAG AGTTAGCTGA TCAAGTTAAA TCAGCAGTGA    | 11040 |
| 25 | CGSAATTTGA AAATGTACAA CCTACAAATG ATCAATTAAC AGATTACAA GAAGCGCATT     | 11100 |
|    | TTGTTGTTTT TGAAAGTGAA GAAAATAGTG AGTCAGTTAT GGACGGCTTT GTTGAACATC    | 11160 |
| 30 | CATTCTATAC AGCAACTTTA AATGGTCAAA AATATGTAGT GATGAAAACA AAGGATGACA    | 11220 |
|    | GTACTCGGAA AGATTTAATT GTAGAAGGTA AACGTGTCAC TACTGTTCTT AAAGATCCTA    | 11280 |
|    | AAAATAATTC TAGAACGCTG ATTTTCCCAT ATATACCTGA CAAAGCAGTT TACAATTGCA    | 11340 |
| 35 | TTGTTAAAGT CGTTGTGGCA AACATTGGTT ATGAAGGTCA ATATCATGTC AGAATTATAA    | 11400 |
|    | ATCAGGATAT CAATACAAAA GATGATGATA CATCACAAAA TAACACGAGT GAACCGCTAA    | 11460 |
|    | ATGTACAAAC AGGACAAGAA GGTAAAGTTG CTGATACAGA TGTAGCTGAA AATAGCAGCA    | 11520 |
| 40 | CTGCAACAAA TCCTAAAGAT GCGTCTGATA AAGCAGATGT GATAGAACCA GAGTCTGACG    | 11580 |
|    | TGGTTAAAGA TGCTGATAAT AATATTGATA AAGATGTGCA ACATGATGTT GATCATTTAT    | 11640 |
|    | CCGATATGTC GGATAATAAT CACTTCGATA AATATGATTT AAAAGAAATG GATACTCAAA    | 11700 |
| 45 | TTGCCAAAGA TACTGTAGA AATGTGGATA AAGATGCCGA TAATAGCGTT GGTATGTCAT     | 11760 |
|    | CTAATGTCGA TACTGTAAAA GACTCTAATA AAAATAAAGA CAAAGTCATA CAGCTGAATC    | 11820 |
| 50 | ATATTGCCGA TAAAAATAAT CATACTGGAA AAGCAGCAAA GCTTGACGTA GTGAAACAAA    | 11880 |
|    | ATTATAATAA TACAGACAAA GTTACTGACA AAAAAACAAC TGAACATCTG CCGAGTGATA    | 11940 |
|    | TTCATAAAAC TGTAGATAAA ACAGTGAAAA CAAAAGAAAA AGCCGCGACA CCATCGAAAG    | 12000 |

55

|    |  |            |            |            |             |            |      |
|----|--|------------|------------|------------|-------------|------------|------|
|    | CATGGTGGGG                                 | CTATTATGCG | TTATTAGGTA | TGTTACGCTT | ATTATTCAGAA | 12120      |      |
|    | AAGAACTCTAA                                | ATAATTANCT | AAATATAGCA | TATGTATGAT | TAACCTTGTA  | 12173      |      |
| 5  | (2) INFORMATION FOR SEQ ID NO: 311:        |            |            |            |             |            |      |
|    | (i) SEQUENCE CHARACTERISTICS:              |            |            |            |             |            |      |
|    | (A) LENGTH: 1316 base pairs                |            |            |            |             |            |      |
|    | (B) TYPE: nucleic acid                     |            |            |            |             |            |      |
| 10 | (C) STRANDEDNESS: double                   |            |            |            |             |            |      |
|    | (D) TOPOLOGY: linear                       |            |            |            |             |            |      |
| 15 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 311: |            |            |            |             |            |      |
|    | CAACATTAAT                                 | ATTGATATTA | AATCTTCTG  | GATAACGTGC | TTGTTTCGAGT | GATAAGTATG | 60   |
|    | CACGCACCTG                                 | ACTTAACTCT | TTATCTAAG  | TAATCGTATG | TTGCTTAGAG  | CTTGTAAT   | 120  |
| 20 | TGCGCTGTA                                  | AAATAACTC  | AATTCATAA  | TGAACCTCG  | TGCGTTTCG   | CTATTATTC  | 180  |
|    | TAACATAAGC                                 | TGAGATCGT  | TTAATTGAAT | TGAAGAAAA  | ATGGGACATC  | ACTTGTGCT  | 240  |
|    | GTAATGACTT                                 | AATCTCAGCA | TCTTCAATA  | ACTTACTTTG | CGTTTCGGCT  | TCACCAAGTT | 300  |
| 25 | CAATTTGGCT                                 | ACTAAAAAT  | TTTGCAATC  | CTTCTGCAAG | TTGACGTTCC  | ACAAAAGTTA | 360  |
|    | AATCATTAGG                                 | TTTGTAAAA  | TACATCTTCA | ATGTACCGAC | GATAGAACCA  | TGCATCTCAA | 420  |
|    | GTGGTATCAC                                 | GATAGCTGCT | CTAAGCGGGC | AATTCGGATG | ACTACAACCA  | ATCTCTCTTT | 480  |
| 30 | TAGTATGAAC                                 | TTCTTTCAAC | TTTCTGATT  | TCAATACATC | TTTACACAGA  | CTTGTAAATA | 540  |
|    | TTTCATTGTG                                 | TGATATGTGA | TGATCACTAC | CTGCACCTAC | ATGCGATAAG  | ATTTCAATTT | 600  |
| 35 | TGCTTGTAA                                  | TGCTACGGCA | GATACCTTCA | TTAAATTTT  | AATAATCATC  | GCAATTTGCT | 660  |
|    | GTGCGGATTC                                 | TCTATTCAAT | CTCTCTTAA  | AATACGGCAA | TGCTCGGTT   | ATCAATTGCA | 720  |
|    | GTAATCATG                                  | TGTTTGAACA | GCCTTCATTT | GCTCTCTTG  | CTTTAATGTT  | GAAATGATAA | 780  |
| 40 | TAGACATAAA                                 | ATTCGCGTGA | CCAACGCTAT | TAACAATAAT | CATTGGTAGT  | GCAATTAATG | 840  |
|    | ATATGAGGTC                                 | AACCGCATAT | GCTTTTGCTG | GGGAAATGT  | TTAAATGCTC  | AACATTTGAA | 900  |
|    | TCATTTCCAT                                 | ACAATTTCCA | ATCATGGCAC | TTTTCGCAAT | ACTCGGGTAA  | CGCTTGGGTC | 960  |
| 45 | TTTGAGCTTG                                 | TAAACAAAA  | TAAACGACAA | TTATACCAAT | AAATATAGAT  | GAGATAAGAT | 1020 |
|    | AAACTTTGTC                                 | ATCCGCCCCA | CCCATATACA | CTCTGAAAT  | ACCTGAAATA  | ACGCCACAAA | 1080 |
|    | ATAGACCTAC                                 | AAAAGGGGCA | CCAACTAACT | CTCGCACACC | TATCGTTAAT  | ACACGTGGTG | 1140 |
| 50 | TAGCTAAAGA                                 | TACATCATCA | TCTAAACGGA | AGTACACACT | TCTGACAAAA  | CTATGTTGAT | 1200 |
|    | GATCGATGAC                                 | GATACAGTT  | AAATTAGACA | TTAAGGCAAA | CAAACTGAAA  | ATAATACATA | 1260 |

## (2) INFORMATION FOR SEQ ID NO: 312:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 7972 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 312:

TATAAATATT ATTTTATTAT CGTTTATCGC TATTTGTGTT ACATAATCAA AACCATAAAT 60  
 TCTTACTCAT TCAGATTAC CCAATATTTT TACTTTTATA ATGTAATGCG TTTTATCCAA 120  
 GTTATTTTTT AAAAATAAAT ATTGAATThG GGGCTGhTTT CATGTCATTA AGAGATGAAG 180  
 CATTGGAAAT GCACAAACGT AATCAAGGTA AATTAGAAGT TAAACCAAAT GTAAAAGTTA 240  
 CTAATAAAGA GGAATTAAGT TTAGCATACT CACCTGGCGT TGCTGAACCG TGTAAGATA 300  
 TTTATGAAGA TAAAAGAAAA GTATATGATT ACACAATTAA AGGAAATACA GTTGCAAGTTA 360  
 TTAATGATGG AACAGCGGTA TTAGGTTTAG GTAACATTGG ACCTGAAGCA AGTATTCTCTG 420  
 TAATGGAAGG TAAAGCAGTA TTATTCAAAA GCTTCGCTGG TATCAATGGG GTGCCTATTG 480  
 CGTTAAATAC AACTGATACC GAAGAAATCA TTAACACAGT TAAGTTGTTA GAACCTAATT 540  
 ATGGTGTAT TAATTTAGAG GATATTTGCG CACCACGTtG TTTTGAAATT GAAGAACGAT 600  
 TGAAAAAAGA AACTAATATT CCGGTATTCC ATGACGATCA ACATGATACA GCAATTGTAA 660  
 CATTGCGCAGG TTTGGTAAAT GCATTGAGAG TTGTTAACAA AGATATTGCT AAAATAAAAG 720  
 TTGTACTAAA TGGTGCTGGT GCAGCAGGAA TAGCCATTGT TAAATTACTA TACGCGTATG 780  
 GTGTAAGAAA TATGCTTATG TGTGACTCAA GAGCGCAAT TTTTGAAGGA CGTTCATATG 840  
 GTATGAATCC TACGAAAGAT GTTGTAGCAA AATGGACAAA TAAAGATAAG ATTGAAGGGT 900  
 CTTTAGAAGA AGTCGTAAAA GACGCAGATG TATTTATCGG GGTTCCTGTA GCTAATGCGC 960  
 TGTACAAGA TATGGTTAAG AGTATGGCAG ATAATCCAAT TATATTGCA ATGGCTAATC 1020  
 CAAATCCTGA AATAATACCT GATGATGCCA AAGCGGCAGG TGCACGAGTT GTTGGTACAG 1080  
 GACGTTGAGA CTATCCTAAC CAAATTAAATA ATGTATTAGC TTTCCCTGGT ATTTTTAGAG 1140  
 GTGCATTAGA GGTGGAAGCT ACACATATAA ATGAAGAAAT GAAAAGGCA GCTGTAGAAG 1200  
 CGATTGCTGA TTTAATCGAT AGTTCTGAAT TAAATGAAGA CTACTGTATC CCAGGACCGT 1260  
 TTGATAAACG TGTAGGCCA TCAGTTGCTC GTAATGTGTC TAAAGCGCA ATGGAATCTG 1320  
 GAGTAGCTAG GATTGAAGTT GATCCGCAAG ATGTGTATGA TAAACAATG AAACCTACAG 1380

|    |   |      |
|----|---|------|
|    | ATTAATGA TGAAATGAA AGTTTATGAT AAACATTCAA CAGTCAAACG AATATAAATC    | 1500 |
| 5  | AAATAAATTT AAACCCGTTT TTAAGTGGTC AAGTTCAGTT TAAGGCTCTA AATGGTTAGA | 1560 |
|    | ACAGAGGTTA TTGGAGGTT TTCTATGTT TAAAGATTTT TTTAATCGAA CAAAGAAAAA   | 1620 |
|    | GAAATATCTT ACAGTACAAG ACTCTAAAAA TAATGATGTG CCTGCAGGTA TTATGACTAA | 1680 |
| 10 | GTGTCCAAAG TGTAAGAAAA TTATGTACAC AAAAGAATTA GCTGAAATTT TAAATGTGTG | 1740 |
|    | CTTTAATTGT GATCATCATA TTGCTTTAAT TGCATATAAA CGTATAGAAG CAATTTCTGA | 1800 |
|    | TGAAGGATCA TTTACAGAA TCGATAAGGG AATGACCTCT GCGAATCCAT TAGATTTTCC  | 1860 |
| 15 | AAGTTATTTA GAAAAAATTG AAAAGGACCA ACAAAGACA GGTCTTAAAG AAGCAGTTGT  | 1920 |
|    | GACTGGTACA GCACAACTAG ATGGTATGAA ATTTGGCGTT GCTGTCAATG ATTCAAGTTT | 1980 |
|    | TAGAAATGGG AGTATGGGAT CGGTTATCGG TGAAGAGATA TGTGTCATCA TTGATTACTG | 2040 |
| 20 | CAGTGAAGAC CGTTTACCCT TTATTTCTTT CTCTGCAAGT GGTGGTGCAC GTATGCAAGA | 2100 |
|    | AGGTATTATT TCCTTGATGC AAATGGGTAA AACCAGTGTA TCTTTAAAAA GTCAATCTGA | 2160 |
|    | CGCTGGACTA TTATATATAT CATATTTAAC ACATCCAAC TCTGGTGGTG TATCTGCAAG  | 2220 |
| 25 | TTTTGCATCA GTTGGTGATA TAAATTTAAG TGAGCCAAAA GCGTTGATAG GTTTTGCAAG | 2280 |
|    | TCGTGAGTTT ATTTGAACGA CAATAAACGA AAAATTGCCA GATGATTTC AACTGCAGA   | 2340 |
| 30 | ATTTTATTTA GAGCATGGAC AATTGGATAA AGTTGTACAT CGTAATGATA TGGCTCAAC  | 2400 |
|    | ATGTGCTGAA ATTCTAAAAA TCCATCAAGA GGTGACTAAA TAATGTTAGA TTTTGAAAAA | 2460 |
|    | CCACTTTTTG AAATTCGAAA TAAATTTGAA TCTTTAAAG AATCTCAAGA TAAAAATGAT  | 2520 |
| 35 | TGGGATTTAC AAGAAGAAAT TGACATGCTT GAAGCGTAT TGAACGAGA AACTAAAAAA   | 2580 |
|    | ATATATACAA ATCTAAACC ATGGGATCGT GTGCAAAATG CCGTTTGCA AGAAGACCT    | 2640 |
|    | ACGACCTAG ATTAATTCC ATATATCTTT GATTCGTTTA TGAACACTCA TGGTGATCGT   | 2700 |
| 40 | AATTTTAGAG ATGATCCAGC AATGATTGGT GGTATTGGCT TTTTAAATGG TCGTGTCTGT | 2760 |
|    | ACAGTTATTG GACAACAACG TGAAGAGAT ACAAAGATA ATATTTATCG AAATTTTGGT   | 2820 |
|    | ATGGCGCATC CAGAAGTTA TCGAAAAGCA TTACGTTTAA TGAACAACG TGAAAAAATC   | 2880 |
| 45 | AATCGTCTTA TCTTTACATT TATAGATACA AAAGTGCAT ATCTGGTAA AGCTGCTGAA   | 2940 |
|    | GAACGTGGAC AAAGTGAATC TATCGCAACA AATTTGATTG AGATGGCTTC ATTAAGAGTA | 3000 |
| 50 | CCAGTTATTG CGATTGTGAT TGGTGAAGGT GCGCTGGAG GTGCTCTAGG TATTGGTATT  | 3060 |
|    | GCCAATAAAG TATTGATGTT AGAGAATAGT ACTTACTCTG TTATATCTCC TGAAGGTGCA | 3120 |
|    | GCGCATTAT TATGAAAGA CAGTAATTGG GCTAAAATG CAGCTGAAAC AATGAAAAAT    | 3180 |

55

|    |             |             |             |             |             |            |      |
|----|-------------|-------------|-------------|-------------|-------------|------------|------|
|    | TAAAGATATT  | GACCATGCTA  | AAAAAGAAGT  | TAAATGTGAT  | ATTTTAAACT  | CTGTGTAGCT | 5100 |
|    | TAAAAACAAA  | AAAGGTGGTA  | ACTTACCTGG  | CGTAGAGTA   | AGTTTACCTG  | GTATTACAGA | 5160 |
| 5  | AAAAGATGCT  | GAAGATATCC  | GTTTCGGTAT  | TAAAGAAAAT  | GTTGACTTCA  | TTGCAGCAAG | 5220 |
|    | TTTCGTACGT  | CGTCTAGTG   | ATGTTTTAGA  | AATTCGTGAA  | ATTTTAGAAG  | AACAAAAAGC | 5280 |
|    | TAACATTTCA  | GTATTCCTCA  | AAATTGAAAA  | CCAAGAAAGT  | ATTGATAATA  | TTGCGGAAAT | 5340 |
| 10 | TCTTGAAGTG  | TCTGATGGTT  | TAATGGTTGC  | ACGTGGTGAC  | ATGGGTGTTT  | AAATTCCACC | 5400 |
|    | TGAAAAAGTA  | CCAAATGGTC  | AAATGCAAGT  | ATGTAACAAAT | TGTAACAAAT  | TAGACAAACC | 5460 |
| 15 | AGTTTATACA  | GCTACACAAA  | TGTTAGATTG  | TATGCAACGT  | AACCCACGTG  | CTACAGCTGC | 5520 |
|    | AGAAGCTAGT  | GACGTGCCA   | ACGCAATCTA  | TGATGGTACA  | GATGCAGTAA  | TGTTATCTGG | 5580 |
|    | TGAACTGCT   | GCTGGTTTAT  | ATCTCGAAGA  | AGCTGTTAAA  | ACAATGAGAA  | ATATTGCTGT | 5640 |
| 20 | ATCAGCTGAA  | GCAGCCCAAG  | ATTACAAAA   | GTTATTGTCA  | GATCGTACTA  | AATTAGTTGA | 5700 |
|    | AAC TTCATTA | GTGAATGCTA  | TCGGTAITTC  | GGTTGTCACAT | CACAGCTTAA  | ACTTAAATGT | 5760 |
|    | TAAAGCAATT  | GTAGCTGCTA  | CTGAAGGTGG  | TTCAACGGCA  | CGTACTATCT  | CCAAATATCG | 5820 |
| 25 | TCCACATTCA  | GACATTATTG  | CGGTGACTCC  | AAGTGAAGAA  | ACTGCACGTC  | AATGTTCAAT | 5880 |
|    | TGTTTGGGGA  | GTTCAACCTG  | TAGTTAAAAA  | AGGACGTAAG  | AGTACAGATG  | CATTGTTAAA | 5940 |
|    | CAATGCAGTT  | GCAACCACTG  | TTGAAACTGG  | TAGAGTATCT  | AATGGTGATT  | TAATCATTAT | 6000 |
| 30 | TACTGCTGGT  | GTACCACTGT  | GTAACCTGG   | ACTATCATAT  | AGATGAAAA   | TCCACATAGT | 6060 |
|    | TGCTGACGAA  | ATTGCTTAATG | GTCACAGTAT  | TGGACGTGGA  | TCAGTTGGTG  | GTACTAGGTT | 6120 |
|    | AGTTGCTGAA  | ACTGTTAAAG  | ATTTAGAAGG  | TAAAGATTTA  | TCTGACAAG   | TTATCGTTAC | 6180 |
|    | TAACTCAATC  | GATGAACCGT  | TGTGACCTTA  | TGTAGAAAAA  | GCTTTAGGCT  | TAATTACAGA | 6240 |
|    | AGAAATAGGT  | ATTACATCAC  | CAAGTGCAAT  | TGTTGGTTTA  | GA AAAAAGTA | TTCCAACAGT | 6300 |
| 40 | TGTAGGTGTA  | GA AAAAGCTG | TTAAAAACAT  | AAGCAATAAC  | ATGTTAGTTA  | CGATTGATGC | 6360 |
|    | TGCTCAAGGT  | AAAATCTTTG  | AAGGATATGC  | AAACGTACTA  | TAATTATATA  | AAAAACGCTC | 6420 |
|    | TTCCATTTAT  | CAACATGGGA  | AAGGCGTTTT  | TGTGGTCATC  | TGGTATTTTA  | TGACGTAATT | 6480 |
| 45 | AATAGGTTAT  | TGTGATATGA  | TAGTGATAGA  | ATGGCAATCT  | ATATAAATGT  | TTATATCTTT | 6540 |
|    | TATACATGTA  | CATTATCACC  | TTCAAACTTT  | CACTCATATT  | ACTTTGAAAA  | TTTATTATAA | 6600 |
|    | AATAGAAGTA  | TGGATGTATT  | TCTGAAATGA  | TACATTATGA  | AATAGATGAG  | AAAGTAAAAA | 6660 |
| 50 | TTTTGAGCCA  | AGTACGCAAT  | TTAATATTAT  | AGAGTGCATA  | TAAACACGGA  | TGGGACATAA | 6720 |
|    | ATCCCTAAAA  | AAACAGCAGT  | AAGATATAATT | TCAAATTAGAA | AATATCTTAC  | TGCTGTTCTC | 6780 |

|    |             |            |             |             |             |             |      |
|----|-------------|------------|-------------|-------------|-------------|-------------|------|
| 5  | CTTCGCAGTG  | GCATGCTCC  | CTCAGGAGTC  | TGCGCATTAA  | TACTACGTAT  | TAACATGTAA  | 6900 |
|    | TTTTACTTTT  | ACATACTTTA | AAAAAATAAA  | ACACTTTGCC  | CAACTTGCCAC | ATAAATGTAA  | 6960 |
|    | AATTCAATAA  | AATGAATTTT | CTGTGTTGGG  | TCCCTTCGTA  | TAATTTAATA  | AATACCACTA  | 7020 |
| 10 | AACATAATTA  | ACGAGGTGCC | TTATGTATAA  | AAITTTATAAC | ATGACCCAAC  | TTACACTACC  | 7080 |
|    | AATAGAAACC  | TCGTGTAGAA | TTCCCTCAAA  | TGATAATTGC  | CGATATGTTA  | ATGAAATTGT  | 7140 |
|    | TGAAACGATA  | CCTGATAGCG | AATTCGATGA  | ATTCCAGACAT | CATCGTGGCG  | CAACATCCTA  | 7200 |
| 15 | TCATCCAAAA  | ATGATGTTAA | AAATCATCTT  | ATATGCATAT  | ACTCAATCTG  | TATTTTCTGG  | 7260 |
|    | TCGTAGAATA  | GAATAATTAC | TTCATGACAG  | TATTCGAATG  | ATGTGGTTAG  | CTCAAGATCA  | 7320 |
|    | AACACCTTCT  | TATAAAACTA | TTAATCGTTT  | TAGAGIGAAT  | CTTAATACTG  | ATGCGTTAAT  | 7380 |
| 20 | TGAATCTTTA  | TTTATTCAGT | TCCATAGTCA  | ATGTTTAAAG  | CAAAATCTTA  | TTGATAATAA  | 7440 |
|    | TTCAATTTTT  | ATTGATGGTA | CAAAAGTAGA  | AGCTAATGCC  | AATAGATATA  | CATTTGTGTG  | 7500 |
|    | GAAGAAAAAGT | ATTCAAAATC | ACGAATCGAA  | ATTGAACGAA  | AATTCAAAAA  | CATTATATCG  | 7560 |
| 25 | TGACTTAGTT  | GAAGAAAAAA | TAATACCAGA  | GATAAAGAG   | GATGGAGATA  | GCGATTTAAC  | 7620 |
|    | AATAGAAGAA  | ATAGATTTAA | TGTGGTAGTCA | TTTAGATAAA  | GAAATCGAAG  | ATTTAAATCA  | 7680 |
|    | TTCTATTGAG  | AACGAAGATT | GTGCTCAAAT  | TAGAAAACAG  | ACCCGTAATA  | AAATACTGTA  | 7740 |
| 30 | GATTAAAGAG  | TTCAAAAAGA | AAITTTGATAT | TTATTCGCAA  | AGAAAAATAA  | ATATGAAGA   | 7800 |
|    | ACAAAAATCT  | ATTCCTTAAG | ATAGACAAATG | TTTTTCTAAA  | ACTGATCTGA  | TCAATGATGCA | 7860 |
|    | ACTTTTATGA  | GAATGAAGGA | AGACCATATG  | AAAATAGGCC  | ACCTTAGCGC  | AGGATACAAT  | 7920 |
| 35 | TTACAAATAG  | CGACAAATTC | TCAAAAATGT  | TTTATCTCAT  | GACCTATTTC  | AA          | 7972 |

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 3175 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

|    |   |     |
|----|---|-----|
|    | ATTTTGTAGT TAATTGCTT TCCTAAAATA ATTTTAGCTT TCATTAAAIT AAACAATTTT  | 60  |
| 50 | ACAAGCTTGG AACACCAATC AAAATCCTAA GTTCTAAAAT GCAATATTAG TAGTCGTTGA | 120 |
|    | CTGAATGAAC ATATGCTTAT AATATTTTTT TGCAATGCTA GTCAAGTTGA TTTTAGCTCA | 180 |
|    | CAAGGATATG CGATTATAT TTCTTACAA CAATGAAAT GCCTGATACA ATGCGATCCT    | 240 |





|    |  |            |            |            |            |             |      |
|----|--|------------|------------|------------|------------|-------------|------|
|    | GTTGATTAC                                  | CGTTTAAGAT | GCCTTTAATG | TTTTTAGCAA | CACCTTCACC | TTGTGCAIT   | 2160 |
|    | GCAATTTGTG                                 | CTGTAGTGG  | TAATGGACGT | TCTTCTCCAG | CTGGGATAAA | CGCTGAACAG  | 2220 |
| 5  | TCACCAATAA                                 | CAAAATGTT  | GTCGTAACCA | TTGATTGTTA | AATCTTGCTT | TGTAACGATA  | 2280 |
|    | CGTCCAGGTT                                 | TAACGCCTTC | AAATGATTCT | TCCATTAAAT | TACTACCACG | TACACCAGCT  | 2340 |
|    | GCCCATACTG                                 | AAGTACCTGC | ATTTAATTGT | TGTTTTTCAC | CATCTACTTC | AACTACAAAA  | 2400 |
| 10 | CCTTTTTCGT                                 | TACAAGCAAC | GATTGGTGTA | GCAATTTTAA | ATTCAACACC | GCGGTCTTCT  | 2460 |
|    | AAGTAGCTAA                                 | CTGCGTGGTT | AACTAATTCT | TCTGAGAAC  | TTGGTAACAT | TTTAGGTGCT  | 2520 |
|    | GCTTCAACAC                                 | AAGTGATTTT | AACTTTATTT | TGATCCACAC | CATATTGCTT | ACATAATTCA  | 2580 |
| 15 | GGAATTCGT                                  | CTGTTAATTC | ACCTAAGAAT | TCAACACCAG | TGAATCCAGC | ACCACCAACT  | 2640 |
|    | AAGATAGATA                                 | AATCGTTATC | ATCTTTTCTT | TTTGATGCTG | CATAGTTAGC | AAATTTGTCT  | 2700 |
| 20 | TGATATGAC                                  | GTGATAATTC | ACGTGCTGTG | ATAACATTTT | CAATTTGGAA | AGCATGATCT  | 2760 |
|    | TTTCATACCT                                 | CGATGCCGAA | TGTTTCACTA | ACGAAACCTA | ATGCTACTAC | TAAATATCA   | 2820 |
|    | AAGTCATAAA                                 | TACCTTGATT | TGTTTCTACC | TTTTTAGCAT | CACGGTCAAT | TTTTGTCTACT | 2880 |
| 25 | TCTGCTGAA                                  | CAAAGTTCAC | TTTGCTTTTC | TTCAAGACAC | TTTCCACAGG | ATATAATACA  | 2940 |
|    | TCTTCATAGT                                 | TTAGTGTAAC | TGCTGATGCT | TCATGTAACC | ATGTGCTTTC | ATAGTGATAT  | 3000 |
|    | TCATTTTAT                                  | TAATAAGCGT | AATTTCTGCT | TCTTCTGTTG | ATATCGCTTT | TTGCAATTTA  | 3060 |
| 30 | GTTACAGTTT                                 | GTAAACCTGC | ATAACCAGCA | CCAAGTACAA | GTACTTTTTT | ACGATCTTGA  | 3120 |
|    | GCCATTTAAT                                 | ThCACTAAG  | ChTTCATATT | TTTTTAACCA | AATGCTGATA | ATTAC       | 3175 |
| 35 | (2) INFORMATION FOR SEQ ID NO: 314:        |            |            |            |            |             |      |
|    | (i) SEQUENCE CHARACTERISTICS:              |            |            |            |            |             |      |
|    | (A) LENGTH: 702 base pairs                 |            |            |            |            |             |      |
|    | (B) TYPE: nucleic acid                     |            |            |            |            |             |      |
|    | (C) STRANDEDNESS: double                   |            |            |            |            |             |      |
| 40 | (D) TOPOLOGY: linear                       |            |            |            |            |             |      |
|    | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 314: |            |            |            |            |             |      |
| 45 | CATCATTATT                                 | AAAGATTTTC | AATCAATACA | gAATCACAAT | ACGTACGCAT | TGTGCACGAT  | 60   |
|    | AAAAATACAG                                 | ATGTGTATAT | TAACATAGAA | CTACAAGAGC | AACTAACGAA | CAAAGCTTAC  | 120  |
|    | ATTGGTGATC                                 | ATATTTATGT | TGAAGGGATA | TGGCTCGAAG | TACAAGCTGA | TGGTTTAAAT  | 180  |
| 50 | GTATTGATGC                                 | AGAATACAGT | GGCATCGTCA | TTAATTCGCT | TAAACAAGA  | GATGCCACAT  | 240  |
|    | GCACAGGCAG                                 | ATGATTACAA | TACGTACCAT | CGTTGCGCAA | GGATTATTCA | CCGTGAACCG  | 300  |

55

EP 0 786 519 A2

|    |            |            |             |             |            |            |     |
|----|------------|------------|-------------|-------------|------------|------------|-----|
|    | TGGCGTTCCA | TTATACGCCC | ATTAGTAATG  | ATTGCTTTAA  | CTGTGTGCAT | CTTTTTAGTG | 420 |
|    | AGACCAATGT | GTTATTATAT | TTTATAGATG  | AITGGATATG  | GTACAGTAAC | GATAGATATT | 480 |
| 5  | GGTATTACAA | CGTATTCTC  | TGAAAAGAAA  | AAGTATAACA  | AAGATGTGTA | AAAACGAGAG | 540 |
|    | AAAGATTACA | AAGCTTATTT | GGATAATAAA  | TCTAAAGAAA  | TTAATAAAGC | GATTAAGAGC | 600 |
|    | CAACGTTTTA | GTTTGAATTA | CCATTATAACA | ACCGTCTGCT  | AAATTAAGAA | TATCGTTGAA | 660 |
| 10 | ACGAAAGCAC | CAAGAATAAT | TGAAAAAACC  | ATCGCGATACA | TC         |            | 702 |

(2) INFORMATION FOR SEQ ID NO: 315:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4121 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 315:

|    |            |            |            |            |             |             |      |
|----|------------|------------|------------|------------|-------------|-------------|------|
| 25 | TATGTTCCGA | CACGAAACA  | AAGTGTAAIT | ACAAGAGCAA | AGATAACTTT  | GAATGTTTGT  | 60   |
|    | AAACGTCCAT | CTTTACCTTC | AGTTAAATGC | ATGAACATA  | ATAATTGARg  | TCCTGCTTGG  | 120  |
| 30 | ACGAATGCAA | AGCCAAAGAT | AATTGTCAAC | TTCCGGTGGa | ATGTTAATGA  | CGTGTATAGT  | 180  |
|    | GTTACGTAAA | CTGCTAAAG  | CGTTAATACG | ATAGATGCGA | TAAATCCTAC  | AGATGATTTT  | 240  |
|    | ATTATTGTAC | TCATCCGCTA | TACACCATCC | CTATCATATA | TACGGCAGTA  | AAGATGAAAA  | 300  |
|    | CCCAAAACA  | ATCTAAGAAG | TGCCAGTATA | AACTTACTAT | AAATAATTTT  | GGCGCATTAT  | 360  |
| 35 | ATTTGCTTAA | TCGCGCTCGT | TGGAAITGGA | TTAATAACAA | AATGGCCCCA  | ACGATACCTA  | 420  |
|    | GCGATACGTG | ACAACCATGC | GTTCCTAATA | GGATAAGAA  | ACTAGACCAG  | TAAAGAACCA  | 480  |
| 40 | TTGTTGGGTT | AAGCGCTTCT | GATGCATAGT | GTGCGAATTC | ATAAAITTCG  | AATCCAACAA  | 540  |
|    | AGACTAAACC | TAAGAATGAC | GTAATGATCA | TCCAAACAT  | CATTAACTTT  | TGTTTTTCTT  | 600  |
|    | GGCGCATGTA | GTAATAGCAA | ATACCACATG | TGTAAGAACT | GAATAATAAT  | GCAAAACGTCA | 660  |
| 45 | TTATTAACAA | AAGAGGCAAT | TCAATAAAT  | CAGTAGTCAT | TTTACCTGCA  | TAATCGCCAC  | 720  |
|    | CATGTTGCAA | AGTTAATAGT | GTGCAAAATA | GGGTACGGAA | TAAAGCAAAT  | TCGGCTGTAA  | 780  |
| 50 | TGAAATCCAA | AAAGCCAGAC | TTATTTAAIT | CGCCTTCATG | TGTCGGTGA   | TCAATAGTGT  | 840  |
|    | TTGTATAGCT | ACTCATAGCT | TACAGCGCTC | CTTTCTTTAA | TGTCAGCTTC  | TCTTAATCTA  | 900  |
|    | GCTTCAGTTT | CTGCAACTCT | AGCAGCAGGG | ATATGATATC | CGTGAATCGAT | TTGAAACATG  | 960  |
|    | CGATAAATCA | TAGTACCAAA | AATACCGAAT | AAACAATAAT | ATGCTGGAAT  | TACAGTTTGG  | 1020 |

|    |            |            |            |            |             |             |      |
|----|------------|------------|------------|------------|-------------|-------------|------|
|    | TTGTTTGSCA | TATGAATGTC | TTTGTAATTA | TGTTGTGCTA | AGTAATGACG  | ACCATGTTCT  | 1140 |
| 5  | TCATATACAA | CAAATGTGTC | GTAGTCATTC | CAATCTGGTG | TAATGGCAAA  | GTTGTATTTA  | 1200 |
|    | GGTGAATTG  | CTGATGCAGT | AGTCCACTCT | AGAGTACGAC | CAAGGCCATC  | CCAGTTATCT  | 1260 |
|    | CCAGTTGCTT | CACGTGGAGA | TTTGAAGTGA | CTGTATACGA | TACTAACAC   | AAGGAATAAG  | 1320 |
| 10 | AATCCGATTG | CCATTAAATA | TGCACCGATA | GTTGAAATTA | AGTTTAATAA  | GAAACCAACA  | 1380 |
|    | TCGTAGTGCA | TATAAGTGTG | TCCAGCAGCT | GGCATACCAT | CTAATCCCAAG | AATGAATTTGT | 1440 |
|    | GGTAGAGAAC | AAACGTTAAA | TCAGTCATG  | AAGCACCAGA | AGCACCATTT  | GTTTAATGTT  | 1500 |
| 15 | TCGTTTAATT | TGTAACCCAT | CATCTTTGGA | TACGACAAGA | TAAACACAG   | TGACAGGCA   | 1560 |
|    | AATACAACAC | CAGTAACCAA | TGTATAGTGG | AAGTGAGCTA | CTAAGAAGTA  | CGTATTGTGA  | 1620 |
|    | TATTGATAGT | CAGCTGATGC | CATTGCTAAC | ATTACACCCG | TAACACCACC  | TAAAGGAAG   | 1680 |
| 20 | TTAGGGATAA | ATGCTAATGA | GAATAGCATT | GGTGACTCAA | ATGTAATACG  | TCCTTTATAT  | 1740 |
|    | AATGTTAATA | ACCAGTTAAA | CAATTTCA   | CCAGTTGAA  | TACCAATCAG  | CATTGTTGAA  | 1800 |
|    | ATTGAGAAGA | ATGAGTTGAT | TAACGCACCA | TTACCCATTG | TGAAGAAATG  | GTAACCCCAA  | 1860 |
| 25 | ACTAAGAAAC | TAAGGAACGC | GATACCGGCA | GTTGCCCAT  | CCATACTTTG  | ATGTCGGAAT  | 1920 |
|    | AAACGCTTAC | GAGCGAATGT | CGGGATAATT | TCTGAGTAAA | TACCAATATG  | TGGAAGGATA  | 1980 |
|    | ACGATATAAA | CTTCAGGGTG | CCCCCATACC | CAGAAGAATG | TAGCCCCAAG  | CATTGGCATA  | 2040 |
| 30 | CGGCCATGTG | CAACTGTGAA | GAATGCTGTG | TCAAATATTC | TATCAGTTGT  | CTATAATGCT  | 2100 |
|    | AACGCTACTG | TTAAGGAGG  | GAAAGCAAGA | ATAACAAAT  | ATGTAGTATT  | AAATGTGTT   | 2160 |
|    | ACTGTAAACA | TGGCATTG   | CATAAATTC  | ATAGTTGGTG | TTTTACATCT  | TAAAAATGTT  | 2220 |
| 35 | ACAAAGAAGT | TGATACCTGT | AGCTAAGGTA | CCAAGCCCTG | AAATTTGTAT  | AGCTATTAA   | 2280 |
|    | TAATAGTTAA | CACCCGACC  | AGGACTGAAT | TCACCTGCTA | GTGGGCGATA  | GTTTGTCCAA  | 2340 |
| 40 | CCAGCTGCTG | GTAACACCAC | AATAATAAAT | GACAGTTGA  | ATAAATCAT   | ACCTGCAAG   | 2400 |
|    | AATAGCCAGA | AACCTAGCTT | GTTTAATACT | GGGAATGCAA | CATCACGTGC  | TCCAAATTTG  | 2460 |
|    | AATGGAACAA | CGATATTCCA | TAAACCAAAG | ATAAATGGCA | TGCCCATGAA  | GATAATCATG  | 2520 |
| 45 | ATTACACCAT | GTGTACTAAA | AATTTCTGTA | TAGTGGTTAG | ATTCTAAAAA  | TTTGTATACA  | 2580 |
|    | GGTACTGTTA | ATTGCGCACG | AATAAGTAAC | GCATCAATAC | CACCACGGAC  | GAACATTAAT  | 2640 |
|    | ACGGCACAGA | TAAATACATA | AATACCGATT | TTCTTATGTT | CTACAGATGT  | GAAACCATCT  | 2700 |
| 50 | TTGTAAGAT  | ATTTCCTATA | TTTAAAGTAA | GTAATCTAGT | CGATTAAAGC  | AATAACTAAG  | 2760 |
|    | AATGGGGCAC | CAATTTGTGC | CATTGTAATC | ATCCAGTTAT | CTTTAACTAG  | TAATGTATCC  | 2820 |

EP 0 786 519 A2

|    |            |             |            |            |            |            |      |
|----|------------|-------------|------------|------------|------------|------------|------|
|    | TTGAAATTTT | CTTCATTTC   | TTGCGATTTT | TCGATTCATC | TTTCTTGAAC | TCATTGTTAT | 2940 |
|    | ATGGTTCGTC | ATTTCCAAGA  | ATCATCAACT | TCATACCATG | TCGTTTATAG | TTCGCAITTG | 3000 |
| 5  | TAATTTGAGC | TTTACGAGCA  | GGTATTAATG | GTTTGTCGA  | TACATCTTTA | AACATATTTT | 3060 |
|    | CTTCACTAGT | GAAGTTTGGG  | TCTTTCAATT | CGAAATTGAA | ACGTTTATAT | GCATAGAAGA | 3120 |
|    | TGTATTCTGG | ATCGGCTGCT  | GGATCAACAA | ACGCCATATG | TGTACCATT  | AATTCTAAG  | 3180 |
| 10 | CTTTATTAGG | TGTGCTTGCT  | ATAATTTGTT | TATCAAAATG | ATCTTGATCT | AACGTTTTCT | 3240 |
|    | TACCTTTAAC | TTCTTTCCAC  | CATTTGTCGT | AGTCTTTTTG | ACTAACGGCA | TTTACTTTAA | 3300 |
|    | ATGTTTGAAG | TGGAATCCT   | TCACCATTTG | AGTTAGAGTT | ACGACCTGCT | AACGTCCAG  | 3360 |
| 15 | TTTGAGATGC | TTCTAAAGTC  | CAATTCATG  | TCATGCCAGT | CATGGCATAT | TTTGACCAC  | 3420 |
|    | CTAATTGTGG | AATCCAGAAA  | CTTGTCATTG | TATCCATAGC | TTGAAGCTTA | AATACAACAG | 3480 |
|    | GACGATCTTT | AGGGATTGTT  | AATGTATTAA | CAGTCTCTAT | ATGTTCATCT | GGATAAGCAA | 3540 |
| 20 | AGAACCATT  | GTATCCTGCA  | CTTACTGCAT | ATACAACCAT | TGGATCTTTC | TCACTCTTGG | 3600 |
|    | GTGGTTTTTC | GTAATCGTAT  | AAAGTTTTAA | CTGTAGGAAT | AGctAAAGCA | GCAACGATTA | 3660 |
| 25 | TGATAGGTAT | TACAAACCAT  | ATTGTTTCAA | TGATGCCAAT | ATGGTGCATC | TTACCAGATT | 3720 |
|    | CGGATCTCT  | ATTATAACTA  | TACTTGTAAA | TAAAAATGCG | GAACTAGCCA | AGTACAACGA | 3780 |
|    | AACAAATAAC | AAGCATGAAG  | ACGATTGAAT | AAAGAATCAA | GAACTTCTGA | CTACTTGCTA | 3840 |
| 30 | CTGGCCCTTT | TGCGTTGAAA  | ATTTCTATAT | TTGAACAACC | ACTAAGTAAA | ATTAGTGTGC | 3900 |
|    | CAATAATAG  | AAGCAAGAC   | TAAATTTTGG | ACACTTTTTT | GACCTCTTAA | TACTACAAT  | 3960 |
|    | TAGGGCTTA  | ACTATAATTT  | TAAATTATTA | CACAATATTT | ACAAGGGCTT | ATGGGAAAAA | 4020 |
| 35 | AATTAATAAA | ATTTTGATCAA | AAATGTGTAT | AAATCAAGGT | GTGACGTGGG | TTCACACATT | 4080 |
|    | TGTTAAAAAT | ATGTGTACAT  | TTTGTGACTA | ATAGCGTTTT | T          |            | 4121 |

(2) INFORMATION FOR SEQ ID NO: 316:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 9310 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO: 316:

50 CGAGTGAGTA CAmACATATT TTTATTTGCA AGGGGTAAAT GGCATATAAC TATCTTTTTT 60  
ATGTAAGCTG GTATAAAATT TATACTAATA GGAGGGATAG TATGAATATA GTAGGGCATC 120

EP 0 786 519 A2

|    |  |      |
|----|--|------|
|    | TTGATTACG ATTAGTGA AAGTCGGTTA ATCAAGACAA TCCTTCAATG TATCAITTTGT    | 240  |
|    | TTTATGGGA CGAAGTAGGT ACAGCCGGAA CAATTTTAAG CTTTTTTGAA ATTCCCAATG   | 300  |
| 5  | CGGGTCATAA GCAGCCAGGT ACTGAAACGA TTTATCGATT TTCATTATTA GTACCAAAATC | 360  |
|    | AAGCGGCACT TCATTATTTT GAAAAACGTC TTGAGAATAA TGGTATTAA TCTGAACGTT   | 420  |
|    | TGTACTATCT TGGACAAGAA GGTGTGTCT TTAAGATGA AGACGACTTA GAAATCATAT    | 480  |
| 10 | TGCTTGTTAA TGATAGTTTT GAAGTACCAC ATCAATGGCA ACATAACGCT TATAGTGA    | 540  |
|    | TACCTCAAGC ATATCAAATT TTAGGAATAG GGCCAGTCGA ATTAAGAGTT AGAAATGCAG  | 600  |
|    | CGCGTACGGT AGAATTTTIG GAAATGTCT TAGGTTATCG CAAAAGAGAT AATAATCAT    | 660  |
| 15 | TCGATGTGCT GACATTAGCA CCACAAGGTT TATATTGGA TTTTGTAGTT ATTGAGCAAC   | 720  |
|    | AGGGACAACG TGAAGACCT GGACGAGGTT ATATCCATCA TATTGCAGTT AATACACCAC   | 780  |
| 20 | AAATGAGTGA CTTAGATGCA ATTTACAAGA AATTACAACA ACAACCACAA AGTAATTCAG  | 840  |
|    | GTATAATTGA TCGCTATTTT TTTAAATCAT TATACTATCG CCATAATTCA ATTATGTATG  | 900  |
|    | AATTTGCGAC TGAAGCGCCT GGATTTACTA TTGATACACC TGTGGAACAA TTGAAGATC   | 960  |
| 25 | AATTGAACTT GCCTGACTTT TTAGAAGCAG AACGTGAACA AATTGAAAGT AAGTTACAG   | 1020 |
|    | AAATATAAAG GAGAATGTTT AATGGCCAAA TTAGAAATGA ATAAAAATAC GCCTCTTGA   | 1080 |
|    | TTTGGTTTGT ATTCTTAGG TGATCATTTA TTGAATCCAT TGAAAGGTGA AAAAGTTAGT   | 1140 |
| 30 | TATGAGCAAC GTATTAATGA AATTATTGAA GCAAGTAAAT TAGCAGATGA AGCAGGTATT  | 1200 |
|    | GATGTTTTTG CAGTTGGTGA AAGTCATCAG GAGCATTTTA CAACACAGGC ACATACGGTT  | 1260 |
|    | GTGTTTAGGT CAATTGCCCA AGCGACAAAG CATATTAAAG TTTCAAGTTC TTCAAGGATT  | 1320 |
| 35 | ATTAGTGCAc AGATCCTGTA AGAGTATTTG AAGACTTCGC GACATTAGAT TTGATTTCTC  | 1380 |
|    | ATGGTAGAGC CGAAATTGTA GCTGGCAGAG CATCAAGAAC AGGTATTTTT GACTTGTTTG  | 1440 |
| 40 | GCTATGATTT AAAAGACTAT GATGAATTGT TTGAAGAAAA ATTAGGTTTA CTTTTAGAGT  | 1500 |
|    | TAAATAAAAC TGAGCGTATT ACTTGGTCTG GAAATATGCG TCCAGAACTT AGAAATATGA  | 1560 |
|    | AAATATTCCC AAGACCAATC GATAATATAT TGCCCAATATG CGGTGCTGTT GGTGTCCAC  | 1620 |
| 45 | CTGCAAGTGC TATTAAAGCG GAAAAACAAG GTGTGCCAAT GATGATTACA ACCCTTGGTG  | 1680 |
|    | GCCACGCAAT GAACITTTAA GGTCTTATAG ATGCTTATCG TCAAGCGGCA ACTGAAGCAG  | 1740 |
|    | GTTTCGATGc TTCGCTTAAG TCTTTACCAG TAAGTACAGC GAGTCTGTTT TATACAGCTG  | 1800 |
| 50 | AAACAACCTCA GGATGCTATG AGAGAATTTT ATCCACATTT GAATACAGGG ATGCTATTTA | 1860 |
|    | TCGTGGTGT TGGTTATCCG AAACAGCAAT TTGCTAATTC GTCAGATTAT CGAGAAGCGC   | 1920 |

55

|    |             |            |            |             |             |             |      |
|----|-------------|------------|------------|-------------|-------------|-------------|------|
|    | GTACATCAACG | TTTTATGGCA | CAGCTTGATT | TTGGCGGTGT  | GCCATTTGAA  | AATGTTATGA  | 2040 |
|    | AGAAATATTGA | GTTAATTGGC | AACGACATTA | TACCGCGGAT  | TAAAAAGCAT  | TTATCAAAAT  | 2100 |
| 5  | AGGAGGGGCG  | TCATCATGAA | TATTGTATTA | TTGTCAAGTT  | CCACAGTAGG  | TTCTAAAACG  | 2160 |
|    | AGAATTGCTA  | TGGATGATTT | AAAAAATGAA | CTAGAAGTCA  | TCAATTGAGG  | ACATCAAAATA | 2220 |
|    | GAGTTGATGG  | ATTTACGAGA | ACTTGAATTA | GAATTTAGCG  | TTGGAAGAA   | TTATCTAGAT  | 2280 |
| 10 | ACTACAGGAG  | ATGTATATA  | ATTAACGACG | TCGTTAATCG  | AGGCTGATGT  | GATTTTATT   | 2340 |
|    | GGTTTTTCAA  | TTTCTCAAGC | TTCCATCCCT | GGTGCTTTGA  | AAATGTGGTT  | TGATCTACTT  | 2400 |
|    | CCAGTCAATG  | CGTTTCTGTA | CAAGGTAAAT | GGACTTTGAG  | CGACAGCAGG  | TTCTAGTAAA  | 2460 |
| 15 | CATTATTTAA  | TTCTCGAAAT | GCATTTAAAA | CCAATATTGA  | GTTACATGAA  | AGCACATACG  | 2520 |
|    | ATGCAACCGT  | ATGTATTTAT | TGAAGAGAAA | GATTTTTCAA  | ATCAACAAAT  | TGTCAATGAT  | 2580 |
| 20 | GATGTTGTAT  | TTCCGTTAAA | AGCGTTGGCA | CAATCCACAA  | TGCGAATCGC  | CAAAGTACAA  | 2640 |
|    | CAACAAGTGT  | TTGAGAAGA  | AAACAACCAA | TACGACTTTT  | AAAGTATAAA  | AATAAGACGC  | 2700 |
|    | TCGGCACACT  | AAATTTGTAA | GTGTTTGAGC | GTCTTTTCAT  | ATTAACATATA | TAGCCAAATGA | 2760 |
| 25 | ACGACGATAA  | AGGCAAGTGA | TGACAAACAT | ATTGAGGTAA  | TAATGATTGT  | CATAAGCGGT  | 2820 |
|    | TTAAGTGCGC  | GATTTTAAAG | ATCTTTAAAT | GCAACATTTA  | ACCCTAAAGC  | AACCATGGCC  | 2880 |
|    | ATTAATAAGC  | AAATGTTTGA | TACAGTATTT | AAAAATTTA   | GCAATGCTGA  | CGGAATGCTT  | 2940 |
| 30 | ACATATCTAT  | TCACATAGGC | CATAATGACA | AATCCATAATA | AAAAGATATG  | AAGTCTTTAT  | 3000 |
|    | CGACCTTTCG  | TAGATGATTC | TGATGAACGG | AAACGCCATA  | TTAAATTAAG  | TACGATGGTT  | 3060 |
|    | AATGGAATCA  | GTAAGAATAC | TCTACCAAGT | TTACCAAGAA  | GTGCAATTTT  | AAGTGATCAT  | 3120 |
| 35 | CTACCACCAA  | AGGCCACCAG | TAAGACAACG | TGTGCAATTT  | CATGAAGACT  | AACACCAGAC  | 3180 |
|    | CAAGCGCCAT  | AAACATTTGT | CGTCATTGAA | AAGATAGCGT  | AGATAGCTGT  | ATATATAAGT  | 3240 |
|    | GAAAATATCG  | TACCAATCAA | TGCGATGATA | CCGATACTAA  | TAGCTGTATC  | CTTTTCCAGT  | 3300 |
| 40 | GATTTGAATA  | TTGGAGCGAC | TGCGGCAATA | CGACGACGAC  | CACAAAGGCC  | TGTGCCGACA  | 3360 |
|    | CCTAGTAATA  | ATGCGATGTT | TTTGTCCACA | TGCAACAGTT  | TGTTGACAAA  | GAGCATCAT   | 3420 |
| 45 | ACAATACTGA  | AAATAACGAC | ACCTACATCG | ATGGCTAATA  | GTTTACTACC  | TTGACCGGATA | 3480 |
|    | ATATCGAATA  | TATTGAGTTT | AAGTCCATAT | AGGATGATTG  | CAAACTCTAA  | TAAATATTTA  | 3540 |
|    | GATGAAAACG  | TAACTACCTG | GCTATATTGT | TCAGGATATC  | CTCTAAAGTG  | ACGATATAGA  | 3600 |
| 50 | ATAGCGATTA  | ATATCGGCAT | AGCTAATGCG | CCAACCTTAT  | CTGAGATGCG  | CAATTTAGCT  | 3660 |
|    | GCTAAAAGAG  | TAAATAATGC | GACTATAAAT | GTTAGTGATA  | GCCCAATCAT  | AAAAAGCTTA  | 3720 |

1171

EP 0 786 519 A2

|    |  |      |
|----|--|------|
|    | ATTTTAAAT ATAAATTGG AATGAATAAT AAAGTAGTGA TTAATAAAG TTGTGTGATA     | 3840 |
|    | GGAACTTGG ACATCAATCA AAGTAATAGG CACTACAACG CTTATTGGCG GGGCCCCAAC   | 3900 |
| 5  | AAAGAAGCTG ACGAAAAGTC agCTTGCAT AATGTGCAAG TTGGGGATGG GCCCAACAT    | 3960 |
|    | AGAGAAATTG GGTCCGTAAT TTCTACAGAC AATGCAAGTT GCGGGGGCCC CAACATAGAG  | 4020 |
|    | AATTTCCGAA AGAAATTCTA CAAGCAATGC AAGTTGGGA AGGACAACAA ATTTAAGATA   | 4080 |
| 10 | CAATGCGTAA CATTAAATATG TTATTATAAT GATAATTTAC AGAATTATAT GAAAAATGAA | 4140 |
|    | TGAGGATGTG ATGGTATGTT TGGAAATGAAA GTGAATGAAC AAATAACATT AAAAAATTTA | 4200 |
|    | GAAGCTCATG ACACAGAAGC GCTTTTCAAT TTAGTCAATC GTTCAAGAAA TTCACTTAGG  | 4260 |
| 15 | GAATGGTTAC CTTGGGTAGA TGCAACTGAG CAACCATCAG ATACGCGTGC ATTTATTAAA  | 4320 |
|    | AGAGGACTTT TGCAATTGTC TGATGGTAAT GGATTTCACT GTGGCAATTT GTATGAAGGA  | 4380 |
| 20 | ACGCTAGTTG GTGTATCGG TTTACATGAA ATTAATCACA TGACACAGAA AACTTCATTA   | 4440 |
|    | GGGTACTATT TAGATAAAGA ATTTGAGGGT CATGGGATTA TGACACAAGC AGTTGAGGCA  | 4500 |
|    | TTGATAAAGT ATTGTTTCGA AGAGCTTGAC TTAACCCGAA TTGAGATTAG TGCCGCGATT  | 4560 |
| 25 | AATAATGAAA AAAGCCGGGC TATTCTTGAA AGGCTGGGAT TTACTAGAGA AGGTATGTTA  | 4620 |
|    | CGTGACAATG AATTACTAAA TGGTATTTAT TCATCGAGTT ACATCTATAG TTTATTAATA  | 4680 |
|    | TCAGATACG ACCAAAAATG ACAAATTAGA CTTACAAAAG AGTGATGACA TTTAAAAATG   | 4740 |
| 30 | CAGCGCTCTT TTATTTAATT TTTGAAAAA AAAGGTTGTT GACAGTATTA TTTTATAACA   | 4800 |
|    | ATATAATGAT TTGTATAATT ATTATCAACT AGATGATGTT TATGGGAGGA TGCTTTAAAA  | 4860 |
|    | CAGCCGTTTT AAGTGTAATG TATTATTTTA GCGTGTAGGG AATGCGAAAA TAATATTTAT  | 4920 |
| 35 | AAGAACACAT CTATGGGGAT AATAGAATTT CTATAATGAG GTGTCAAAAT GAAAAAGTTA  | 4980 |
|    | ACAACGCTAT TATTAGCATC AACGTTATTA ATTGCTGCAT GTGGGAACGA CGATAGTAAG  | 5040 |
| 40 | AAGGATGATT CAAAGACATC GAAAAAAGAT GATGGTGTTA AAGCAGAATT AAAACAAGCA  | 5100 |
|    | ACAAAAGCAT ATGATAAATA TACTGATGAA CAGTTAAATG AATTTTTAAA AGGTACAGAA  | 5160 |
|    | AAATTTGTTA AAGCGATTGA AAATAATGAT ATGGCCCAAG CAAAAGCGTT ATATCCAAAA  | 5220 |
| 45 | GTTCGTATGT ATTATGAACG CTCTGAACCA GTTGCAAGAG CATTGGAGA TTTAGATCCT   | 5280 |
|    | AAAATTGATG CACGCTTTGC AGATATGAAA GAAGAGAAAA AGGAAAAAGA ATGGTCAGGA  | 5340 |
|    | TATCATGAAG TTGAAAAAGC ATTATACGAA GATAAGAAAA TTGATGATGT GACTAAAAAA  | 5400 |
| 50 | GATGCACAAC AATTATTGAA AGATGCAAAA GAATTGCATG CCAAAGCTGA TACATTAGAT  | 5460 |
|    | ATCACACCAA AATTAATGTT ACAAGGTTCT GTTGACCTAT TAAATGAAGT TGCAACTTCT  | 5520 |

55

|    |  |      |
|----|--|------|
|    | GTTGAAGGCG CACAAAAAT TTATGACTTA TTAAACCTA TTTTAGAGAA AAAAGATAAA    | 5640 |
|    | AAATTAAAGT ATGATATCCA AATGAACCTC GATAAAGTGA ATCAATTATT GGATAAATAT  | 5700 |
| 5  | AAAGATAACA ACGCCGGTTA TGAGTCATTT GAAAAAGTAT CGAAGAAAGA CCGTAAAGCA  | 5760 |
|    | TTTGCGGATG CTGTTAATGC ATTAGGAGAG CCACTAAGTA AAATGGCTGT GATTACTGAA  | 5820 |
|    | TGACAAATTA TGAACAAGTT AACGATAGTA CGCAATTTTC AAGACGTACA TTTTGA AAAA | 5880 |
| 10 | TGTTAGGTAT TGGCGGTGCC GGTGTTGCAA TTGGCGCAAG TGGTGTGGT AGCATGTGGT   | 5940 |
|    | CTTTCAAATC AATGTTCAAT ACACCAGAAG ATCCGGA AAA AGATGCGTAT GAATTTTATG | 6000 |
|    | GTAAGTGCA ACCAGGCATT ACCACACCCA CGCAAAAAAC ATGCAATTTT GTTGGCTTAG   | 6060 |
| 15 | ATTGTGAAGT AAAAGATAGA GATGCAATTA AGGCAATGTT TAAAAAGTGG ACGGTTATGG  | 6120 |
|    | CTGATCOTAT GATGGATGGT GATACAGTTG GCAAGCCGAG TAACAATCCT TTAATGCCAC  | 6180 |
|    | CAGTAGATAC CGTGGAATCG ATAGGATTAG GTGCAAGCAA GTTAACGATT ACCTTTGGGA  | 6240 |
| 20 | TTAGTAAGTC TTTGATGAAG AAAATTGGGT TATCTAGTAA AATTCGCCAT GCCTTTAAAG  | 6300 |
|    | ATTTACCGCA TTTTCCGAAT GATCAGTTAA TAGACGATTA CAGCGATGGT GATATTATGA  | 6360 |
| 25 | TTCAAGCATG CTCAAATGAT TCGCAAGTAT CCTTTCATGC GGTTCATAAT TTAGTTCGTC  | 6420 |
|    | CATTCGAGA TATTGTTAAG GTACGTTGGG CGCAATCTGG TTTTATCTCT GCTAAAGSTA   | 6480 |
|    | AGGAAACACC TAGAAATTTA ATGGCATTTA AAGATGGAAC AATTAATCCT AGGAAGAATA  | 6540 |
| 30 | ATCAACTTAA AGATTATGTG TTTATTGATG ACGGATGGGC GAAACATGGA ACTTATTGCG  | 6600 |
|    | TTGTGAGACG TATTCAAATA CACATTGAAA CGTGGGATCG TACTGCGCTG GAAGACAAG   | 6660 |
|    | AGGCTACATT TGGTCGGAAG CGACATAGTG GTGCACCGTT AACAGGTGGG AAAGAGTTTG  | 6720 |
| 35 | ATGAAATTGA CTTAAAGGCG AAAGATAGTC ATGGCGAGTA TATTATTGAT AAAGATGCCC  | 6780 |
|    | ATACGAGGCT AGCGAAAGAA GCAAAATACGT CAATTTTACG TAGAGCCTTT AATTATGTGG | 6840 |
| 40 | ATGGTACGGA TGACCGCACA GGTAACCTCG AAACAGGCTT ACTTTTATT GCTTTTCAAA   | 6900 |
|    | AAGCGACAAA ACAATTATTC GATATACAAA ATAATTTAGG TAGTAATGAT AAATTAATG   | 6960 |
|    | AATATATTAC ACATAGAGGT TCTGCTTCAT TTTTAGTATT ACCAGGTGTT AGTAAGGGAG  | 7020 |
| 45 | GATACCTTGG TGAACCATTA TTTGACTAAA TTTGTAGCAA TGCTAATAAC TGCTGCTATG  | 7080 |
|    | GTGTGTAGCT TTGGGTTACT GAAAGTCAG GCAGCAGAAC AACAAAGTAT TAGTGATGA    | 7140 |
|    | TATAGTGTGA TAACGGATGC GAAATCTGCA CTTTCTAATA ATTCGATATC GAATGACAAT  | 7200 |
| 50 | AAGCAGAAAG CAATTGAGCA AGTGGTAAGT GCAGTTAAGA AATTATCGCT TGAAGATAAT  | 7260 |
|    | AGTGAAAGTA ATGCTGTCAA ATCAGATGTG AGAAAGCTTG AAGATGCAAA AGCGAATGAT  | 7320 |



EP 0 786 519 A2

|    |  |      |
|----|--|------|
|    | GCTAGTAAAG ATGCGGGTTC TAAATTTAAA CTATTGCAAC AGCAAGTCGA TGCTAAAGAT  | 7440 |
|    | GCTGCGATGA CAAAAGCGAT TAAAGATAAA AATAAAGCGG AATTAGAATC TTTGAACAA   | 7500 |
| 5  | AGTTTGAATC AGATTGGAC AAGTAATGAA ACAGTGATTC GCAATTATGA CGCAATCAA    | 7560 |
|    | TATGGACAAA TTGAAGTCGC ATTATTACAA CTTAGAATTG CAATTATGA GTCAACATTA   | 7620 |
|    | GATACGGCAA AAGTGTGACA TGCTTGGACA ACTTTTAAAT CAAATATTGA TCATGTGCGAT | 7680 |
| 10 | AAAAAAGTA ATACGCTCTGC AAATGATCAA TACCATGTAT CACAATTAAA TGATGCGTTA  | 7740 |
|    | GAGAAGGCGA TTAAGGCTAT CGACGACAAT CAATTGTGCG ATGCTGcaTg TGCCTTACA   | 7800 |
|    | CATTTTATAG AAAGTTGGCC GTATGTTGAA GGTCAAATTC AAATAAAGA CGTGCTTTG    | 7860 |
| 15 | TATACGAAAA TTGAAGATAA AATACCATAT TATCAAAGTG TATTAGACGA ACATAATAAA  | 7920 |
|    | GCACATGTGA AAGATGGTTT AGTAGATTTA AATAACCAA TTAAGAGGT TGTGTGCCAT    | 7980 |
| 20 | AGTATAGCT TGTCGATGT GATGATTATC TTTTACGCTG AAGGGCTAGA AGTGTGTGA     | 8040 |
|    | ATTGTAATGA CATTGACTAC CATGACGCGT AATGTAAAG ATAAGAAAG GACTGCAAGT    | 8100 |
|    | GTGATTGGTG GTGCAATTGC CGGACTTGTA CTGAGTATTA TCTTAGCAAT TAGCTTTGA   | 8160 |
| 25 | GAAGCTTTAG GGAATAGTGG CATTCTTCGT GAAAGTATGG AAGCGGAGT AGGTATCGT    | 8220 |
|    | GCGGTCATAT TAATGTTTAT CGTGGTGTTT TGGATGCACA AACGTTCAAA TGCAAAACGT  | 8280 |
|    | TGGAATGACA TGATTAAAAA TATGTATGCT AATGCGATTA GTAATGGTAA TTTGGTATTG  | 8340 |
| 30 | TTAGCGACGA TTGGTTTAAT ATCTGTGTTG CGTGAAGGTG TCGAGTTAT CATTTCTAT    | 8400 |
|    | ATGGGGATGA TAGGTGAGCT AGCGACCAAA GATTTTATTA TTGTATTGC TTTAGCTATC   | 8460 |
|    | GTTATTTTAA TCATCTTTGC ATTATTATT AGATTATAG TTAATTAAT ACCTATTTTC     | 8520 |
| 35 | TATATATTTA GAGTGTTGTC GATCTTTATT TTTATTATGG GATTCAAAAT GCTTGGCCTA  | 8580 |
|    | AGTATTCAAA AGTTACAATT ATTAGGTGCG ATGCCAAGAC ATGTTATTGA AGGATTCCCA  | 8640 |
| 40 | ACGATTAACT GGTGGGCTT TTATCCAAGT TATGAACCAT TGATAGCACA AGGTGCTTAT   | 8700 |
|    | ATTATGGTAG TTGCTATCTT AATCTTTAAA TTTAAAAAAT AAAAACAAGG CCGAGTGCGT  | 8760 |
|    | GTTTTTTTTT TTGCTATATT GGAATATTTC GGTATTGCAG TATAACGATA ATCACAGCAT  | 8820 |
| 45 | TGATTTCGTAT AAGGTTAATG TGTGGCGGT TTGCTCTGCG ATGTGAACCT AACGATGAAC  | 8880 |
|    | ATACTGAACCT CAAAGAGCAA TATGAGTGGC AATGTGAGTA ATATATTTAA TGTTAAATCG | 8940 |
|    | GGTGTGCAAT TGATACTTGC TAATACAAAG CAAGCGAAAT AAATATATTT ACGTTAATGT  | 9000 |
| 50 | TTCAATGATG TGATATCTAT AAGACCGAAT TTTGCAAGAC CCATAAAATA TATTGGTAAT  | 9060 |
|    | TGAAATAGAA GACCAATGT GAATAACCAA CGTATGAGTT CAATCAAATA TGCTTTAAAG   | 9120 |

55

EP 0 786 519 A2

GGAAAGCCAA CATAAAATGC AAAAGCGACG CCAGCACAGA ATAATAACAC GCTGAAAAAA 9240  
 CTATATTTAT AATAAAATG ACGTTCATTA TTATGTAATC CAGGTCCAAT GAATGCCAC 9300  
 5 AATTGATAAA 9310

(2) INFORMATION FOR SEQ ID NO: 317:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 3458 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 317:

ATTAGTATTA CACACTAATG TAAATAGATT GGTGGAGAA GAGATATTTG CTAATAAGTG 60  
 CCTTGCTAAT AATGATGTTT AAATTATGAA TTCAATAAAA AAATTAATTG AAGCTGAGTT 120  
 ATTAACAACG ACTAATGATT TTGAAGTTAG TATATATAAA AAGACAAGC CTGAATTACA 180  
 AAGTATTTTA AAAAGTTTTG GTATAAAAC AACAGGTAAT AACCTGACT TAATTAAACG 240  
 25 TAITGACGAC AATTTTCATA TTATTAAATA CTTAGATTTA CCATATGTAT ATATACCAAT  
 TAAGAAAGGA GAAGAAATTT TAAAGAAAAC AGAGTACTTA ACCTCTTTTA TACAAAGTTA 360  
 TGGTGAAATT TCTCTTGAGC GTGCTTATTA TTTGGTTGAA AACTATATAG ATGAAAAATTG 420  
 30 CGATGATAAA GTTGACAGAA TATACAAGTT TGAATTTCAA AGAAAAATATG ACAATGGCGA  
 GTTTGATTTT AATCATCGAT ATAATTTTCA ATTGAATATG TTGATAGATC ACTATAAAG 540  
 AGATGTAAAA GACTACGATA ATGCCAGAAA GTATTCAAAT ATTTATCTTT ACTTTGGTTT 600  
 35 GAGAGATTTT TTAATAAAAT TAATGAGTAA TTATTCATAT TACGATAGTA AAGGGAATAT  
 AGATTGTAAC GAAATACAAA ACGATCTGAA TAGATTTATC AACTCTAGCG CTTCGTTAT 720  
 GTACGAGCGA TTAATATATA ATGAAAATTT ATCCAATAAT ATTATGTTT AATTATTATA 780  
 40 AAAGACACA CAAGATTATA GTGATTGGA AGAACAAATG ATTGAAAAGT TCATAAACTA  
 TGTAGTGTCT AATGTAAAAA AAGAAAGTAG GAGTAATACT CTTATAGAGT TATCAAAAAT 900  
 45 TTTGAGAAC GGATATACAA TTGATAAGA AGAATTTAAA AAAGAAGATG ATTATCTTTC  
 TAAGTACATA TTTACTGACA TAGATTATTT GAAAAAGTTA GAATCAAAAA TAAACGTTGC 1020  
 TATTGATATT CGAAGTGGAG AAATTCATT GGTATTAGAT GATGATAGCC TTGATATATT 1080  
 50 AATACAAAAT CAAAAATCG GCAATGAGTT TTGAGTCATA ACTAAATATA ATATGTGTA 1140  
 GAGAGTTTTT ATTTAAATGG CTAANAATTGG TTATGATCGT GTATCAACGA AAGATTAATA 1200

EP 0 786 519 A2

|    |  |      |
|----|--|------|
|    | AACGTACAGA GCTTGTAAAG TGTTTAGATT ATTTACGAGA GAGCGACACA TTAGTTGTCT  | 1320 |
|    | ATCAACTTGA TCGGTTAGGT AGAACGACAA AACTATTaAT TGAATTaTCA CAATGATTCTG | 1380 |
| 5  | ATGATAACGG AATTGACTTA CAAATTAGTA ACATGAACAT TTCAACGAAA GACACAATGG  | 1440 |
|    | GCAAAATGTT TTTTACGATG ATGAGTGCAAT TTTTCaGGAT TAGAAGTTAA TTTACTATGT | 1500 |
|    | CAGTGTATAA AATAGACTTA GCAGCAACAA GAGCGAGAGG CCGAAAAAGC GGGCGCCCTT  | 1560 |
| 10 | TTTTACGAGA GAATAAAAAA TGAGAAATTA AATTTTtATA TGATGAACAA ACGATAACAG  | 1620 |
|    | GGGAAGAAAT AGCTAGTTAG ACAAGGGTAT GTCACTCAAC TGTTTATCGA GTTGATTAAAG | 1680 |
| 15 | AAATGAAAAA ACTTATACTA TGAATTACTG TTTAAAAGTG TGCATGTTAT AATATTTAT   | 1740 |
|    | GAGCAAGTTG GATAGATGGT GGCTAATCTC TTAATAAAGG GGTGATGCCT ATGGTTATAG  | 1800 |
|    | TGTCTACTCC TAGaAAAGGA CTAGCATGTC TGATTTTGAA ATGCTTATGG TTGTATTAAAC | 1860 |
| 20 | AATCATTTGGT TTAGTATTGA TTAGTACTCA AGACCATAAA AAATAaCCTT CTATTCGCTT | 1920 |
|    | TGACCGGCAT TTTTGAAAGC TATTTTtAAA TAAAATATAA GGTCAACGTC TTTTtAACGG  | 1980 |
|    | GCTCATTAGG GTAACATGTT TCCGAGTGTT GCCCTTTTtG TGTTTCAAGA GTTAATGATT  | 2040 |
| 25 | TTCACTCTTT GCTTCTACTT GCTACAAATA TATTTTAACA CATTTTCTT ATGAATTGTA   | 2100 |
|    | GTTCTGAACA TAATCAGAAT TAATAAAACC AACTTTCCAT ACAGCAGAAA ATACAATTAA  | 2160 |
|    | AAGTATAGAA TGTAATCACG AATTATATAA AAAGTATTGG AGACCTCTTC ATATATAGAT  | 2220 |
| 30 | AATTCACCTA GTTATTTTAG AAAGAAGCCC CTAAaCACTA AAGTTGAAAA ATAGAGAAAC  | 2280 |
|    | ACAGTTgGAT TAOGCATCAA CTGCATaAGg CCCCTAAmAA CTAAAGTTGT AAGGGGcYCT  | 2340 |
|    | AAAATTTATT TTGGTTGATT GTCTTCTGGT TTATCIGAAG TCATTGTTTT TGTTGTATTA  | 2400 |
| 35 | TCATTTAATG AATCTGTGTC TTTTCTGTT TTGTGTTCAA CAGATGTTGT CACTTTATCA   | 2460 |
|    | TTTTCTACTT TTGTATTTTT TGTTTCAGTG ACAACTTTTT TATTCTTGTC AGCTGTTTTT  | 2520 |
| 40 | ACTTTTCTCT TAGTAATTTT TTGACCAGT TTTAAGTAAT ATTGAACGAT ACCCATTTAA   | 2580 |
|    | ATGATTGCAT GAATAACAAC GAATAAAATA ATTGTcATTA CTGTATAAAC TCTTATGATA  | 2640 |
|    | TTTTCTGCAA CACTTTGAGA GAGTGATTGC GTATTAAATG AAATCAAGTA ACCAACAGGT  | 2700 |
| 45 | GTTTTAAATA AAATAATAAT TAAGTTTAAT AATAGTATTC CAATGAAAAA TTtGAAAAAT  | 2760 |
|    | GTTTTTTGAC CATTTTTTcAT TGCTTTAAAT CCAITAGCTA AGTGtGTTTT TACTTTATCG | 2820 |
|    | TTAGTTGATT CaACGAACT AGTcAtAAAG TTGAAATAG GTATTAAcAA TAACCAAGTA    | 2880 |
| 50 | ATAATTGCTT TAATCAATAA AACAAcAATG ATTATGATGC TTtGAGTTGT AATAGCAATA  | 2940 |
|    | CCAATCAGGT GTGTACTATC TGCATATGAT GATTGTACTG AATTCATAAT CATTTCTGAT  | 3000 |
| 55 |  |      |



EP 0 786 519 A2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 319:

|    |   |     |
|----|---|-----|
|    | AATATTGATa TTATATATAG TCGAATCaAT GATGACCCaC ATCGACCTAA TGCAAAATAT   | 60  |
| 5  | AAGACAAITCA AGCAGITAAA TGACTTGTAT TGCTGCGGA ATTAAAGATA TAATGAAATC   | 120 |
|    | GCATTCAAGA AATTGTATAG TCATTTGCTA TCAATTTTCAG AAATATATTA TGAATTGCTA  | 180 |
|    | AATTTAATAA AAGCGAGTGA TCAGTATTAG AGAGAATAGA GCGTTAAGAC TCTATCGCCG   | 240 |
| 10 | AAGCGCAAGT AATTATTAC GAACTCTCA GGcAAAAGGr TAATACTGTa ACGCGTTCCT     | 300 |
|    | GAATTGGTGA TTTATAAACA GGTAGCGAT TGCTATCCTG TTTTATAAT TTTAAGGGGG     | 360 |
|    | TATTTCAATG TCAAGTGATT TAAACAAAC ACCTTATAT CAAAATTATG TTGATAGAGG     | 420 |
| 15 | TGCAAAAATT GTGGAATTC GAGGATGGGC GATGCTGTT CAATTTTCAA GTATTAAAGA     | 480 |
|    | GGAGCATAAT GCTGTTTCGAT ACGAAATTGG CCGTGTTCGAT GTTAGTCATA TGGGTGAAAT | 540 |
| 20 | TGAAGTAACA GGTAAAGATG CTAGTCAGTT TGTGCAATAT TTATTATCAA ATGACTACTGA  | 600 |
|    | TAATTTAACT ACTTCAAAAG CATTATATAC TGCTTTATGT AATGAAGAAG GCGGTATTAT   | 660 |
|    | TGATGATTTA GTAATATATA AATTAGCTGA CGACAATTAT TTATTAGTTG TTAATGCTGC   | 720 |
| 25 | TAATACTGAA AAAGATTTTA ATTGGATTTT AAAACACAAA GAGAAATTTG ATGTTGAAGT   | 780 |
|    | ACAAAATGTA TCAAAACCAAT ATGGTCAATT AGCAATACAA GGACCAAAAG CcAGAGATTT  | 840 |
|    | AATTAATCAA TTAGTTGATG AAGAkGTAAC TGAaa                              | 875 |

30 (2) INFORMATION FOR SEQ ID NO: 320:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 5897 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 320:

|    |   |     |
|----|---|-----|
|    | TCTTTTACG AAGGCTACCG TTTAGGATCT GATAACTTAA CTGTAGGAGA AATTGTATT   | 60  |
|    | AATACAGCGA TGACAGGTTA TCAAGAACT ATTTGAGATC CATCATATAC AGGTGAGATC  | 120 |
| 45 | ATTACTTTTA CGTATCCATT AATCGGTAAT TATGGTATCA ATAGAGACGA TTTTGAATCA | 180 |
|    | TTAGTACCTA CATTAAACGG TATTGTAGTG AAAGAAGCGA GTGCGCATCC AAGTAATTTT | 240 |
|    | AGACAGCAAA AGACACTTCA TGACGTTTTA GAATTGCATC AAATTCAGG GATTGCGAGT  | 300 |
| 50 | GTGTATACAA GAAGTATTAC CGGTAAATTT CGACAACACG GTGTGTTAAA AGCTGGTTTT | 360 |
|    | ACTGATCGAA AAGAGATAT TGATCAACTT GTCAACACTT TACAACAAGT AGAATTACCT  | 420 |

55

EP 0 786 519 A2

|    |  |      |
|----|--|------|
|    | AGTGTCTGAC TTGTAGACIT TGGTAAGAAG CAAAATATTG TTCGAGAAIT AACGTGAGA   | 540  |
|    | GGTTGTAAACG TCACAGTTGT ACCATATACA ACTACTGCCG AAGAAATTTT AGCAATGGCT | 600  |
| 5  | CCAGATGGCG TTATGCTATC AAACGGACCA GGTATCCTG AAGTTGTAGA ATGTGCGATT   | 660  |
|    | CCAATGATTC AAGGAATTTT AGGGAAAATT CGGTCTCTTG GTATCTGTCT AGGACATCAA  | 720  |
|    | CTTTTGTGAT TATCTCAAGG AGCAAGCTCA TTAAAAATGA AGTTTGGTCA TCGTGGTGGG  | 780  |
| 10 | AACCAATCCAG TTAATAAATT AGAGACTGGA AAAGTTGATA TTACGAGTCA AAACCATGGA | 840  |
|    | TATGCAATAG ATATAGATTG GTTAAAAAGT ACTGATTTAG AAGTTACTCA TCCTTGCATTA | 900  |
| 15 | AATGATGGTA CTGTAGAAGG TTTAAAAACAT AAAACATTAC CAGCATTTTC TGTTCAATAC | 960  |
|    | CATCCTGAAG CAAATCCAGG ACCGTGAGAT TCAAACATC TATTTGATGA TTTTGTAGCA   | 1020 |
|    | ATGATGACTA ATTTTAAAGG AAAGGAGCGT CATATCAATG CCTAAACGTA ATGATATCAA  | 1080 |
| 20 | AACAATTTTA GTAATAGGGT CTGGGCCAAT TATCATAGGT CAAGCAGCTG AATTTGATTA  | 1140 |
|    | TGCTGGAAAC CAAGCATGTC TAGCTTTAAA AGAAGAGGGA TATCGAGTTA TTCTTGTAAA  | 1200 |
|    | TTCAAATCCA GCGACAATCA TGACTGATAA GGAATTTGCG GATAAAGTAT ATATCGAACC  | 1260 |
| 25 | GTAACTCAT GATTTTATAG CGCGAATTAT ACGTAAAGAG CAACCTGACG CTTTACTTCC   | 1320 |
|    | AACTTTAGGT GGTCAAACAG GTTTAAACAT GGCAGTTCAA CTACACGAAA GTGGTGTGCT  | 1380 |
|    | TCAAGATAAT AACGTCCAAT TATTAGGAAC TGAGCTAACA TCAATTCAC AAGCAGAAGA   | 1440 |
| 30 | CGGTGAAATG TTTAGAACAT TAATGAATGA TTTAAACGTT CCTGTACCAG AGAGTGACAT  | 1500 |
|    | TGTAATATCA GTAGAGCAAG CCTTTAAATT CAAAGAGCAA GTGGGATACC CGCTAATTGT  | 1560 |
|    | TAGACCGGCA TTTACGATGG GTGGTACCGG AGCGGTATT TGTCATAATG ATGAAGAATT   | 1620 |
| 35 | ACATGAAATC GTCTCAAATG GTCTTCATTA TAGTCCAGCA ACGCAATGTT TATTAGAAA   | 1680 |
|    | ATCTATCGCA GGTTTTAAAG AAATCGAATA CGAGTAATGc GTGaTAAAA CGATAATGCC   | 1740 |
| 40 | ATCGTTGTAT GTAACATGGA AAATATTGAT CCAGTTGGTA TTCATACAGG CGATTCAATT  | 1800 |
|    | GTTGTGGCTC CTAGTCAAAC ATTATCAGAT GTTGAGTATC AAATGTTACG TGATGTTTCA  | 1860 |
|    | TTAAAAGTTA TTGAGCTTT AGGTATCGAA GGTGGTTGTA ATGTTCAATT AGCATTAGAT   | 1920 |
| 45 | CCCCATTCAT TCGATTATTA TATTATAGAA GTAAATCCGC GTGTATCAG TTATCAGCG    | 1980 |
|    | TTAgCTTCAA AAGCAACAGG ATATCTTATT GCAAAATTAG CTGCTAAAA CGCGTTGGT    | 2040 |
|    | CTAACATTAG ATGAATGTT AAATCCAATT ACAGGAACAT CTTATGCAGC GTTTGAACCA   | 2100 |
| 50 | ACTTTAGACT ATGTGATTC AAAAATACCA AGATTTCCTT TTGATAAATT TGA AAAAGGA  | 2160 |
|    | GAACGAGAGC TTGGCACACA AATGAAGCA ACAGGTGAAG TTATGGCCAT TGGTCGAAT    | 2220 |
| 55 |  |      |

EP 0 786 519 A2

|    |   |      |
|----|---|------|
|    | TTACCAATG GTGAAAGCTT CGATCTTGAT TATATTAAAG AACGTATTTC ACACCAAGAT  | 2340 |
|    | GATGAACGAT TATTTTTCAT CGGCGAACAA TTAGAAGAGG CACAACATT GAAGAAATTC  | 2400 |
| 5  | ATAATATGAC TCAGATTGAT TACTTCTTCT TACACAAGTT CCAAAACATT ATTGATATTG | 2460 |
|    | AGCATCAATT AAAAGAGCAT CAAGGTGATT TAGAATATCT TAAATATGCA AAAGATTATG | 2520 |
| 10 | GATTTAGTGA TAAAACAATA GGCATCGCT TTAATATGAC GGAAGAAGAA GTATATCAAT  | 2580 |
|    | TGGTATGGA AAATGATATT AAACCTGTT ACAAGATGGT TGATACTTGC GCAGCTGAAT   | 2640 |
|    | TTGAATCTTC AACACCATAT TATTATGGTA CATACGAAAC TGAAAATGAA TCCATAGTTA | 2700 |
| 15 | CTGACAAAGA AAAAATCTTA GTATTAGGCT CTGACCAAT TCGAATCGGC CAAGGTGTAG  | 2760 |
|    | AATTTGACTA TGGACAGTT CACGCCGTTT GGGCAATTCA AAAAGCAGGG TACGAAGCGA  | 2820 |
|    | TAATTGTGAA TAACATCCA GAAACAGTTT CAACAGACTT CTCAATTTCT GACAATTAT   | 2880 |
| 20 | ACTTTGAACC TTTAACTGAA GAAGATGTGA TGAATATCAT TAATTTAGAA AAACCTAAG  | 2940 |
|    | GTGTGTTGT ACAATTTGGA GGACAAACAG CGATTAATTT AGCAGACAAA TTGGCTAAAC  | 3000 |
|    | ATGGTGTTAA AATACTTGGT ACTTCACTAG AAAATCTAAA TCGTGCTGAA GATAGAAAAG | 3060 |
| 25 | AATTTGAAGC ACTATTAGA AAAAATTAACG TGCCACAGCC ACAAGGAAA ACAGCTACAT  | 3120 |
|    | CACCTGAGGA AGCATTAGCG AATGCTGCG AAATCGGATA TCCGGTTGTA GTAAGACCTT  | 3180 |
|    | CTTATGTATT AGGTGGTCGC GCAATGGAAA TTGTAGACAA TGACAAAGAG TTAGAAAAT  | 3240 |
| 30 | ATATGACCCA GGCTGTAAGA GCGAGTCCGG AACATCCGGT ACTAGTCGAT AGATATTTAA | 3300 |
|    | CTGGTAAAGA AATTGAAGTT GATCGGATTT GTGATGGAGA AACGGTCATT ATTCCAGGAA | 3360 |
|    | TCATGGAACA TATTGAACGT GCTGGTGTGC ATAGTGGTGA CTCAATCGCT GTATATCCAC | 3420 |
| 35 | CACAACTTTT GACAGAAGAC GAGTTAGCAA CACTTGAGGA CTATACTATA AAATTAGCTA | 3480 |
|    | AAGGTTTAAA CATCATTGGC TTAATCAACA TTCAATTCGT TATAGCTCAC GATGGTGTGT | 3540 |
| 40 | ATGTTTTAGA AGTAAATCCA CGTCTAGTA GAACGGTACC ATTCTTAAAT AAAATTACTG  | 3600 |
|    | ATATTCCAAT GGCACAATTA GCTATGCGAG CAATCATTTG GGAATAACTA ACAGATATGG | 3660 |
|    | GTTATCAAGA AGGGGTTCAA CCAATGCTG AGGGTGCTT TGTGAAAGCA CCAGTATTTA   | 3720 |
| 45 | GTTTTAATAA ATTGAAAAAT GTTGATATTA CTTTAGGACC TGAAATGAAG TCAACAGGTG | 3780 |
|    | AAGTGATGGG GAAAGATACT ACATTAGAAA AGGCGTTATT CAAAGGGTTA ACAGGTAGTG | 3840 |
|    | GCGTTGAAGT TAAAGATCAC GGTACAGTAT TAATGACCGT CAGTGACAAA GATAAAGAGG | 3900 |
| 50 | AAGTTGTTAA ATTGGCACAA CGCTTAAATG AAGTTGGCTA TAAATTTTAA GCAACGCTG  | 3960 |
|    | GACAGCTAA TAAATTAGCT GAGTATGACA TACCTGCAGA AGTAGTAGGC AAAATTGGTG  | 4020 |

55

55



TTAAAAACAA AAAAGCT

5897

(2) INFORMATION FOR SEQ ID NO: 321:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 7965 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 321:

|    |  |      |
|----|--|------|
| 15 | TCATCTCTAAA TCAACTTATT TTCCATTGCA TAAATTGCTG CTTGTGTACG ATCGCTAACT | 60   |
|    | TGTAATTTCAC TAAATATATG ACTGACATGT GTTTTAATTG TTTTTCAGA TACAAATAAA  | 120  |
|    | GTCTCTGCAA TCTCTTTATT TGTTTTACCT TTAACCATTT CACGTAACAC TTCAATTTCT  | 180  |
| 20 | CTCTTTGACA ACTTAATCGT GTAGTGTGGT TTTTGGCTAA CTGTTTCGAA TACATCTTGT  | 240  |
|    | GCCTTAGGAT GTATCATTTT TTCACGTTTC ATAACCTCGT TAATAGTTTC AATTAATTGC  | 300  |
|    | TGAGGCTCAA CGTCTTTTAT TTCATAACCA TCAGCACCTT TATTGATTGC TGAAATTACA  | 360  |
| 25 | TGTTTCATCAT CAACATAACT TGTTAATACC AAAACTTTAA TATCCGGATA ATGTGCCTTA | 420  |
|    | ATATATTCGG TAATTTCAT ACCATTTCATG CCAGGCATCA CTAAATCTAA TAGCACAATA  | 480  |
|    | TCAGGGTGCT CATGCTCTTT TAAATATTCT AAAAATGTTT TCCTCATCTGC AAAGTCTTGT | 540  |
| 30 | AAAACTCTTA TGTTTTCAAT CGTGGATAAT AAAAAATGCA ATCCTTGTGC CACAATATAA  | 600  |
|    | KGGWCATCTA CTAATAThAC TTTGTTTCATG GGTATCTCC TTAAATCAAG CTATTTTATA  | 660  |
|    | GGAATTGTGA ATTGTATTG TGTACCCCTT GTTGCTGAG AATGAAAGGT CACTTTACCT    | 720  |
| 35 | CTTAATAAAT TAACCTCTTG TTTTATGITA TTAATACCGT GTGATGAAGC TATCTGAACA  | 780  |
|    | TTATCGATCT CAAATCCTTG ACCATAATCA ATCAGTCAA TATATAGTAT ATCGTTCAAT   | 840  |
| 40 | TGTTTTAATG TAAGATCCAT TTTATTGTA TCAGCATGTT TCTTAACATT ATTAATACAC   | 900  |
|    | TCTTGTAATG CTCTGTATAT GTTTTCTTCG ATTTCATTAG ATAAATCGAT TAAACCTTCT  | 960  |
|    | ACATTTCAT TTAATTGTAT ATGCATTAAAT TTACTATATG CTGTCAAAGC ATGAATTAAA  | 1020 |
| 45 | CCTTGCTCAA GTCCAACTGG CTTAAGTnGC CAAATCAATG CACGCAATTC ATTAACGGCA  | 1080 |
|    | TTTTGACTCG TTeCCTCAAT CGTCTTGAAT GCTTGTTTAG CGATGGATTG GTTTGACATG  | 1140 |
|    | CCATACGCAG CATGTGCTGT TAGTTTTACA GAAATTAACA TTTGATTTAC TGAATCATGT  | 1200 |
| 50 | AAATCTCTAG CTAACAGATT ACGTTCAITA ATTTTTGCGG CTTCCTTTTC ACGGTCGTGT  | 1260 |
|    | AAATAAATAC GTTTGATGGC TGACCCTAAT TGAATGCGA CAGACTCTAG CAACCTAAAA   | 1320 |

|    |   |      |
|----|---|------|
|    | TGACCCGATT TAAGTGGTAC CGTTGCATGA TGTGTAATAT TGTCAATTTG GCTAGGAAAT   | 1440 |
|    | GCTTTAGAGG CTAAGTTAAT ACGAGAACAA TTGACGATAT TCGACGCTTT CATTAGCCTA   | 1500 |
| 5  | CGTTGATTAA ATGCTTTCAC ACACCAACAA GACCCATCTT TAATATAGTG ACAGTGGTCT   | 1560 |
|    | GCTGTCAAAG ATTGTGGTAG AGCCACATGT GATACAAGTT CATGTTCAAC TACGCTATTG   | 1620 |
|    | ATGAAAATA TCCAGCTGT CGTGAAATTA CTGCCCTCAA TTAATATTTT TAACGCACCT     | 1680 |
| 10 | TGGTTCATGC TATACATTTT TGTTCCTTCG TTTAAAAATT CGGCAATCTC TTTTAATAAA   | 1740 |
|    | GCTATGCGG TCCTTTGTTT CATCAATTCG CTCCAATTCG TTTTACGTG TATTAACTAT     | 1800 |
|    | TATACATTGA GTTATTATAT TTTTAAATCT TAGACGTAAA CATGATAAAA TGGCCTTGAT   | 1860 |
| 15 | TACTCAATAG TTATATTTCG GAGAACTGAT TTGTGATATG ATATTAAAGA CTATAGGAGG   | 1920 |
|    | ATTTTATGAA ATTTAAAAATA CCAGAAAACCT TTAATGACTT AAGTTTACGA GATATTTTCC | 1980 |
|    | AACAACCTTA GGTACCTAAA AAAGATTAC ATCAATTAAA TATGTCTAAA GATATTACTA    | 2040 |
| 20 | TTAATGATAA ACCTGCGCGA TTAATGGATA AAGTGCATAC TGGCGAGCAT GTATTGTGTC   | 2100 |
|    | CAACCATCGA TGAAAAAAGT AATTATGTTC CAAGTTATCG TTATGCACAA ATTAAATACG   | 2160 |
| 25 | AAGACGATGA TATGGCAATC GTAATGAAAC CTAAGGGTGT TAAGACTCAC CCTAATGATT   | 2220 |
|    | TAAAAAGAAAG CAATACTTTA ATGAATCATG TGATTTCAC TATTGATAGT GACTATGTG    | 2280 |
|    | AACCAATTCG TCGACTOGAC CAGGAAACAG TAGGATTATT AATTGTGCTT AAAAACTCT    | 2340 |
| 30 | TAATGAAAAA AATTCTTGAT CGCATGTTAG AAGACAATGA TATTACGCGG ATATACAAAG   | 2400 |
|    | CAAAATGTTAA GGCACCTTTA CCTTTAAAC CACAAACGAT TGATATGCCA ATTTGGTAAAG  | 2460 |
|    | ATAAAATCCA TTGGAATAAA CGACGTGTGT CTCCTACTGG ACAGCGTGCA ATTACACACA   | 2520 |
| 35 | TTTTAACTTC AAAAAAGATA AAAGAAGCTG TGTGCCACT TGAAATCAAG TTGGTACTG     | 2580 |
|    | GACGTACTCA TCAAAATCGT GTGCATTTAG CTGAAATGG TCACCCGTGT ATTGGTGATC    | 2640 |
| 40 | CTTTATATGG TGATTCAACG TTAAGACAA TAGAACTGA AAGTTACAAA ATAGAGTTTG     | 2700 |
|    | TGATCCCTT GACTAAGGAA GTCATTTCGG TTTCCTTGG TACTAATTT GATTAGTTTT      | 2760 |
|    | GCATGATATT TAAACATGCA ATACCGCATT GTAACTAAAT CAAGTATCAA CTAAACGGA    | 2820 |
| 45 | TAGATGGAAA ATTATTAATT TTTTCAGATG TTCGGTTTTT TTGTTTTTT CGATGCTTAG    | 2880 |
|    | GATTTTATAT TTTGATATTT TAGTAATTAT TCATTTTATA ACATCCTTGG ATAATGACTT   | 2940 |
|    | GTAGTCTTTT TCAACTGCGT TACGTGTATC TATGGACAA ACATGACATC ATAAGATTTT    | 3000 |
| 50 | TATCACAGGT TGTTTGGCCA ATACATGTAC AACAAATCAT CATATAAAA ATAGGTTCTA    | 3060 |
|    | TAATAAAACG GACTCCATGA AAAGTTTTTC CTTTTCATGG CTCTATATCA AATCAGACTT   | 3120 |

|    |             |            |             |             |            |             |      |
|----|-------------|------------|-------------|-------------|------------|-------------|------|
|    | CCAAACCTTTA | TTTTATATTA | ATATTTAATT  | AATGAGGATC  | TACCATATCT | TCTGTGTTAA  | 3240 |
| 5  | TCCATGCTTC  | AAATGTGTCT | TCTGTAACAT  | ATCCAGTTTG  | AATTGCAGAT | TCTTTTAAAG  | 3300 |
|    | TTAAACCTTC  | TTTATGGGCT | TTCTTAGCAA  | TTTGAGCTGC  | TTTTTCATAA | CCAATATGTG  | 3360 |
|    | GATTTAATGC  | AGTAACTAAC | ATTAATGATT  | GATTTAAATA  | ATTATCAATA | TTCTCTTCGA  | 3420 |
| 10 | TTGGTTCAAT  | GCCCACTGCA | CAATTGTTAT  | TAAATGTTTC  | CATACCATAA | GCTAAAGATG  | 3480 |
|    | AAATGTAGTG  | TAGTGTATTA | TGCATAATAA  | CTGGTTTATA  | AACATTCAAT | TCAAAGTTAC  | 3540 |
|    | CTGTGGAact  | TGCGAACCAA | CAACTGTATC  | ATTACCCATT  | ACTTGGAGCT | CTACCAATGT  | 3600 |
| 15 | TAACATTICA  | CATTGTGTAG | GATTAACTTT  | ACCAGGCATA  | ATTGATGAAC | CTGGTTCATT  | 3660 |
|    | TTCAGGGATA  | GAAATTTCTG | CCAAACCAGC  | TCGTGGCCCT  | GAAGCCAACC | ATCTCACATC  | 3720 |
|    | ATTAGCAATT  | TTCATTAAGT | CTCTGCTPAA  | TGCGCTCAAT  | GTTCCATGCA | ATTGAACAAC  | 3780 |
| 20 | TTCATCATGC  | GCTGTAAAGT | CGTGGAAATT  | ATTTTCAGAA  | GATACAAATG | GATAACCCGT  | 3840 |
|    | ATTTTCTGAA  | ATATAATGTG | CCACTTTATC  | ACCAAAITCA  | GGATGCGCAT | TAATACCAGT  | 3900 |
|    | ACCAACAGCC  | GTACCACCGA | TGGCAAGATT  | TAAAAATGTG  | TTCTTAGATT | CAGATAACAT  | 3960 |
| 25 | TGTTTCACAA  | CGGTCAAGCA | TATAACGCCA  | GCCCAATACT  | TCTTTGCTCA | GTTTGATCGG  | 4020 |
|    | CGTTCATCT   | TGTAATATGT | TACGACCAAT  | TTTAAATAAT  | GAATCAAAAT | TATCTTCTKT  | 4080 |
|    | TTCTTTCAAA  | GATTTTCTTA | AAGTTTITTA  | TGCAGGTTCT  | AATTTTGTTT | CAACCTCTGT  | 4140 |
| 30 | ATATAATGCA  | ACGTGCATAG | CAGTTGGGaa  | TGTATCaTt   | GaACTTTGAG | ATTyATTIAC  | 4200 |
|    | ATCATCATtK  | GGGTGGATAC | TTTCATCACT  | TTGATGATCT  | TTTAAATACA | TATTAGCAAC  | 4260 |
| 35 | ATAACTTACT  | ACTTCGTTCA | CATTCATATT  | ACTTTGTGTA  | CGCGTCTCTG | TTTGCCATAC  | 4320 |
|    | AACTAGTGGG  | AAGTGTTTAT | CTAAATTCACC | TGATAAAAAT  | TGATCACATG | CGTATACAAT  | 4380 |
|    | GGCATCTTTC  | TTTGCTCTCG | TAAATTTTCC  | TAAATCAAAA  | TTAGCTATTG | CTGCTGCACG  | 4440 |
| 40 | CTTAGTTTGT  | GCAAAACCAT | AAACTACTTC  | GATTGGCATA  | CGCTCTTTAC | CAACTGGGAA  | 4500 |
|    | ATTACGTTTA  | CTTCTTTCTG | TTTGAGCACC  | CCAATATTTA  | TCTGCAGGTA | CTTCTATTTC  | 4560 |
|    | TCCAAAAGTA  | TCAATGTCAA | TTCTTAATGA  | CATTCACATT  | CTCCGCTTAT | CATGCTTTTAT | 4620 |
| 45 | TTAACTGTAG  | TATATCATTA | AAATAAFTTA  | TTCGACCAAT  | TATGATTAAA | ACGTTTTCAT  | 4680 |
|    | AATTTGAAAT  | AAAAATACAC | TAATCGCACG  | TGTTCAACCCT | TTAATACAGT | GATACGGTCA  | 4740 |
|    | TACGATTAGT  | GTGTTATCTA | TCATTATTTA  | GTTATTATTG  | AACTAAGTTT | AATTACGATA  | 4800 |
| 50 | CTTTGTTTTA  | GTAGCTTCAA | CCGAGCAAT   | AGCTGTAAGT  | ATATATAATA | CAGCACTAAC  | 4860 |
|    | AATTGTGCTA  | TATGGATTTA | GAGCAACAAG  | CGTaCCTAAA  | ACTCCTGTTA | AACTGCGATA  | 4920 |

EP 0 786 519 A2

|    |  |      |
|----|--|------|
|    | TACAATACCT GATTGATTAC TTTTAATGAA TGTTCGCGCA TTAACATCAT CAATTAATCC  | 5040 |
|    | TTTGTATAAA TTGAGTTGTA ATTTTATTAC TTTGAAAATA ACAGGTAAAT ATAATGCCCC  | 5100 |
| 5  | AAITGCCAAT GGAAAAGCTT TAATTGATAT TAAACTTATA ATAACGTGTG CTATCAATAA  | 5160 |
|    | TTGAATCCAG TATTTTCCTA ACATAAATAT ATAAATCTCC TCTAATTCA TTCTCAATA    | 5220 |
| 10 | GCATATCATA ATCTGGCAT ATTAAGAAAC GCGGTTTAAT GATTTCATTA AAAATATTAC   | 5280 |
|    | TGATATGA TA CTCTTTCAA TTATGTCTGG AGTAATTAAT TATCAATTCC GTTTAAATGG  | 5340 |
|    | TGTTTAAATA TTTAAATTG AACTTTGTAT ATATTACTAT GTCTGGTACA CAAATCAATG   | 5400 |
| 15 | TTTATGCTT TACAAAGTTA TATTGGCAGT AGTTGACTGC AGTCCAACAT ATAGAGGCTT   | 5460 |
|    | CGGAATGTCA GCTTCTATTT CATGCAAGTT GGTGGAGCTC CAACATAGTG GAATTGGATT  | 5520 |
|    | CCCAATTCTT ACAGACATTG CAAATTGGGG AAACGGGCCA CAAACTCAGA AACTGGTGA   | 5580 |
| 20 | AAGTCAGCTT AAAATAACAT GCAAGTTGGC GAGGCCCAAA AATAGTGAGA TCGGATTTCT  | 5640 |
|    | AATTCTTACA GACATTGCAA ATTAGGGAAA CGGGCCACAA ACTCAGAAGT TGGTGGAAG   | 5700 |
|    | TCAGCTTAAA ATAACATGCA AGTTGGCGGG GCCACAACAT AGAAAAATTG GATCCTCAAT  | 5760 |
| 25 | TTCTACAAAC AATGTAAGTT GGGGAAACAG CCCCACACT GAAACTAGCA GAAAGTCAGC   | 5820 |
|    | TTCTATGAAT ATAATAAAAA AGCTAGGTAA CAAAATGCTA CCTAACTTCA TATTCAAGAT  | 5880 |
|    | AATCAATCCT ATTTGATATA TGTCTATAC TATACATTAT TTACATGATA AATAACTGAA   | 5940 |
| 30 | TATTACACA TTTATAACT TTTACTGACTG TCTTCTTCAG AATTCTTTTC TTGATCATTT   | 6000 |
|    | TGATCAGAGA TTTGTTCCaT TTCTTTACCT AATTCTTTTA AATCTTCAA ATCCGTTACC   | 6060 |
|    | ATACTGTTTT CTCTTCATG ATAATTAAAT TTTGGATCTT TGTCTTTAGA CATAATCAAT   | 6120 |
| 35 | ACCTCACAGA TTTTAAATA AGCAAAGCGT GACGTAAAGT AAGCTTTAAC ATCTTCAGGT   | 6180 |
|    | AAACCTGCAG CCGCTTCTTT ATCAAGAATA ACATTTACCA TTCTATGTGC TTTTAAATCG  | 6240 |
| 40 | GCTGGTTCGA AGCTTGTCTT ACCATTTTCT TGATATAATT TTTCAACTAC ATCTCGTTTA  | 6300 |
|    | TTAGCACCTG TCACTACTAA GAAAATTCTT CTGTCTTCCA TTAGTCTCTG ACGAATACTA  | 6360 |
|    | ACATTTAACT TACCTTGCTC ATCGATAGAA ACAACTTGTA ATGTTAATTT CCCTTTATTT  | 6420 |
| 45 | TCTTTAGTTT TAATCTTATC AGCGATTAAAT TCGATTGCAT CTTTTTCATA AGCAATTGGA | 6480 |
|    | TAAACTTGAC CTGCTGGTAC ACCTAACGCT TCGAAATATG ATTTTTTATC GTCATAATCT  | 6540 |
|    | AAAATATTTA TTTGGCTAAA ATCAACAGCA TGTTTTTCAA CATTTTCTTT TAATTCACTT  | 6600 |
| 50 | AGAACTGGCG CTGATCTGT ATCTAAATGA AAACCTGCAA TTGTAGTAGG ATTATTGTGA   | 6660 |
|    | AATTGCTTTC TAATAATATC AGCAGCATAT TCTGCTACAA GTTGACTATT GTCAAGACT   | 6720 |
| 55 |  |      |



TGTAGTTGTT GAAAATATTT ATCGACGCTT AACAGATTCA GAAGAACAAC TAAAAGGTGA 300  
 AAATTTAATT ATCAGTGGGA CAACTGAAGT ATTTAAACCA ATAATGTCAT CGACACTAGT 360  
 5 TACTATTATC GTCTTCTTAC CACTTGTGTT TGTATCAGGT TCAGTAGGCG AAATGTTTAG 420  
 ACCTTTTGCA TTGGCTATTG CATTTAGTTT ATTAGCATCG TTATTAGTGT CAATTACACT 480  
 CGTTCAGCG TTGGCAGCTA CACTATTTAA AAAAGGCGTT AAACGTCGTA ATAAACAACA 540  
 10 TCAAGAAGGA TTAGGTGTTG TTAGTACAAC TTATAAAAAA GTATTACATT GGTCATTAAA 600  
 TCATAAGTGG ATTGTAATTA TATTAAGTAC ATTAATTTTG GTTGCAACTA TTGTATTTGG 660  
 AGGACCGAGA CTAGGCACTA GCTTTATTTC AGCAGGTGAC GATAAATTTT TAGCTATTAC 720  
 15 TTATACACCG AAGCCTGGTG AAACGGAGCA AGCAGTGTG AATCATGCGA AAGATGTTGA 780  
 AAAATATTTA AAACAGAAAA AGCATGTAAA AACAAATCAA TACTCAGTTG GCGTAGTAG 840  
 20 TCCAGTAGAT CCAACGGGTA GTACAAATAG TATGGCAATC ATGGTTGAAT ATGATAATGA 900  
 CACGCCATAT TTTGATGTAG AAGCGGATAA GGTATTAAAA CATGCAGATG GCTTTAAACA 960  
 TCCTGGAGAG TGGAAAAATC AAGATTTAGG AACAGGTGCA GGTAAATAAT CTGTAGAGGT 1020  
 25 TACTGTAAAA GGTCCATCAA TGGATGCCAT AAAATCACT GTAAAGATA TTGAACAGAA 1080  
 AATGAAACAG GTTAAAGGAC TAGCCAATGT CAAATCTGAT TTATCGCAAA CATATGATCA 1140  
 GTATGAAATT AAAGTCGATC AAAATAAAGC GGCAGAAAAA GGTATTTCTG CAAGTCAACT 1200  
 30 TGCAATGCAC TTGAATGAAA ACTTACCAGA AAAAACAGTT ACGACTGTTA AAGAAAAATGG 1260  
 TAAAACGTGT GATGTTAAAG TCAACAAAAA TAAGCAAAAC GC 1302

## (2) INFORMATION FOR SEQ ID NO: 323:

35 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1003 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear  
 40

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 323:

45 ATATATATTA ATTTAAACGT GTTTCACATG TACCAGTGT AATGACAGAT AATGCTGCGT 60  
 TTAACACACC TTCAACAAGG TTTTGTACTG CTTCATCAGA GAAGAAAGCA ATATGTGGTG 120  
 TTACTAAAAT TCTTTCATGT TCGATTAACT CTAATAATGT TTTATCGTCA ATGTCCTTAT 180  
 50 TAGTCCAGTC ATTTGTGAAG TATGCTGCTT CATTTTTCATA AGTATCAACT GCAGCACCTA 240  
 ACAAAGTACC ATCGTTCCTT GCAGCGATTA AATCAGGTGT ATTGATGACT GCACCACTGT 300

55

GATAGCTTTC TTTGTCGCT GGAACATGTA AAGAAATAT ATCGCATCT TTAATAGCTT 420  
 CTTTAACTACT ATCTTTATAA GTTAAAAAGT CTAAATCTTT ATTAGGATAG GCGTCATAAG 480  
 5 CTGTAATGT AGCACCAAT CCTGCATATA TtTAGCTGT AGCAGCACCG ATACGACCCG 540  
 TACCAATAAT TGCAACAGCT ATATTTTAA CTGGTTTGA CATGATTCT GCTTGCCAAAG 600  
 10 TAAAATCATG TGCTTGTACA CGGCGTTCAA TATCTGGGAA GCGACGCACT AATTGTAGGG 660  
 CGATAGATAC AGAATACTCT GCAATtGTTT CAGGTGAATA ACTAGGAACG TTAGATATCA 720  
 CAATATTGTG TTTTITAGCT AAATCTAAAT CATACATATC AAATCCAGCA GTACGTTGTG 780  
 15 CAATTGTGTT AATACCGTAA GATTCTAATT TAGGATAAAC GTCATTTTCT AACTTACCAA 840  
 ATTGCAITGT AGTACGCCA TCGTAATCTT TTAATTGATC GACTGTAGCA CTGTGATAATA 900  
 GCTCTTTAGA AGTAGTTACT TCGACATTAT TCTtTtTCCC CCAATTTAAT GCCATCTCTT 960  
 20 TCTCATAATC ACGCGTACCA AAGAACATAA TTCTCGTCAT TAT 1003

(2) INFORMATION FOR SEQ ID NO: 324:

(i) SEQUENCE CHARACTERISTICS:  
 25 (A) LENGTH: 5030 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 324:

GGCTTTTATA GTGTTTACGA TTATTATGAT TACTTATATT CACGCTGTTA AAAATTATCC 60  
 35 AAATAACCGT ACTGTTTATT ATGTTTATAC AGCTGCATTT ATACTTGTTA TATTACAAGT 120  
 TATCACAGGT GCATTATCTA TTATGACAAA TGTTAACCTG ATAATCGCAC TTTTCCATGC 180  
 ATTAITTTATC ACTTATTAT TTGGTATGAC AACATACTTT ATCATGCTAA TGTTACGATC 240  
 40 AGTAAGAAGT GACATGCAAT AACAAAAAAG GmAAACCGTAA TTTTAATGGC ACGCCCATTA 300  
 AAATTACCGT nTTTATATC AATATTTAAA AtTAAACcTA AGCCATGTAA AAACGAGATT 360  
 45 ACACGCTAAT TGTTGTGTAA TCTCGTTTTA TnTTAATCAT TTTAGTCAGT TGCTTTTCA 420  
 ATTTGATTA ATAAATGCC TGTCGCTATT GTGTACCATT TATTACAGT TACTTGTTA 480  
 ATCACACCGT CAAATGGTGC TTGAATTGTT GTTTCCATTT TCATAGCTTC AGTAATTAGC 540  
 50 AACGGCTGAT TAGCTTTCAC AGTTTCACCT ACATAACCT TGACTTCAGT TACTGAACCT 600  
 GGCATTGAG CACCGATATG ACTTGGATTA CTCTTATCTG CTTTGGCTT AACGTTGCA 660  
 TTTGTATGCA CATTTTCATC TTTAATGTAA ATACGTCTCG CTTGACCATT CATCGCATAG 720

55

|    |   |      |
|----|---|------|
|    | CGTTTACCTT TATCGATTTC GATTTCCTACT GTTTCACCAT TACGCATTCC AAAGAAGAAT  | 840  |
|    | GTAGGCGTAT CAAGTAACGA TAAGTTTCCG TATTGATTTC TAGTTTGAAT ATATTGTTCA   | 900  |
| 5  | TATACTTTTG GATATAGTAC ATAACATAA ATATCTTGCT CCGTAACAGG ACCTTGTTGC    | 960  |
|    | TCTTCTTCAA GCAACTCACG GACTTTTTCA AAATCAACTG GCTCTAGATA TTCACCTGGA   | 1020 |
|    | CGAGCTGTTA GTGCTTCTTG GCCTTTTAAA ATAACGGCTT GTAAATCTTT ATTA AAAACA  | 1080 |
| 10 | TTTACAGGTT GTCCATTTTC ACCTTTGAAG AACGACATA CTGATTCTGG GAAATCTAAT    | 1140 |
|    | TTATAGCCAT CTGTAATCAC GGATTGTTCA TCAAGATCAT TTTGTACCAT ATAAAGTGCC   | 1200 |
| 15 | ATATCACCAA CTACTTTAGA CGATGGTGTT ACTTTTACGA TATCACCAA TAAGAAATTC    | 1260 |
|    | ACTCTGCGAT ACATACTTT GACTTCATCA AATCTTTGCG CTAAACCTAA ACTTTTAGCT    | 1320 |
|    | TGTTGACTTA AATTGGAATA CTGTCCACCA GGCATTTCAT GTTGATAAAT TTCAGTATTC   | 1380 |
| 20 | GGTGATTIGA TATCACTTTC AAAGTCTGAA TAATAAGTAC GTACAGTTGA CCAATAATGA   | 1440 |
|    | CTAAGTGACT CCATACCTTC AATATCAGTT CTAAGTGGC GTGGGAAGCC ATTTAATGCA    | 1500 |
|    | TAATATAACG AATTGGCGCT TGGCTGACTT GTTAAACACC TCATTGAAGC AACAGCAGTA   | 1560 |
| 25 | TCAATGATAT CGACACCAGC ATCTATTGCT TGTGTTGATG TTAATAAACC ATTACCACTT   | 1620 |
|    | GTATCATGAG TGTGAAGATG AATTGGTAAA TCTACAGCTG ATTTTAACCT ACCAATCAAT   | 1680 |
| 30 | TCGTAAAGCG CTTTAGGTTT TAATAAGCCT GCCATATCTT TAATCGCTAA AATATGGAAT   | 1740 |
|    | CCTTCACGTT CTAACCTCTT AGCTAGTTTG ACATAATACT CTAAAGTATA AATGTTTGAT   | 1800 |
|    | CGCTCAGGAT TTAATAATGC ACCTGTATAA CAAATAGTAC CTTCAGAGAT TTTGCCCGCT   | 1860 |
| 35 | TCTTGTACTG CTTTCATGGC AACTTTCATT TGATCTACCC AGTTTAATGA ATCGAAAATT   | 1920 |
|    | CTAAAGACAT CTATGCCCTG TTAGCAGCTT TCTTGTAGCA ATTTATGAAT AACATTATCA   | 1980 |
|    | GGATAGTTTT TATAACCAAC TGGCGTTGAA GCACGTAAAC ACATTTGGAA TAATACATTT   | 2040 |
| 40 | GGAATAGCTT TAGCTAGACG TTCAGTCGT TCCCATGGGT TTTCTTCAA GAAATTATAT     | 2100 |
|    | GCCACATCAA ATGTAGCACG GCCCCACATT TCTAGTGAGA AACCATCTTT AAATACGTCC   | 2160 |
| 45 | GCTGTTTGGG ATGCGATATT AATCATATCC TTAGTCTCAA CTCGTGTAGC TAATAATGAT   | 2220 |
|    | TGTTGTGTCAT CTCTAAAGGT TGTAATCTGTT AGTAAGACAT CATCTGCTTT TTTAACCCAT | 2280 |
|    | TCAGCTACAC CTTTGGACC TACTTCATCA AGCAATTGTT TCGTACCACT AAATAGGCG     | 2340 |
| 50 | ATTTTACTTG AAGATACAGT TGAATTGAT GCTAATTCAT AGTCTGGTTT CGGACGTTTC    | 2400 |
|    | TCAACATTTG GGAACCAATT AATTGTTACA TTACCTATAT ATTCTAATGT TTTAGTACCT   | 2460 |
| 55 | CTATCTAGAG ACGGCTGAAT GTCGAAAAGT TCTGGTGTTC CTTCAATAAA TTTAGTTGTG   | 2520 |



|    |  |      |
|----|--|------|
|    | TTAACACCAC GAATACGCAT TTCTCGTAAT GAGCGTACCA TTTTTCCTTC TGCTTGITTA  | 2640 |
|    | AATGATATCG CGTGTGTAGA TAATTTTACG AGTAATGAAT CATAATAAGG TGATATCTCA  | 2700 |
| 5  | GCACCTTGGA AACCATCTCC AGCATCAAGA CGTACACCAA AGCCCCCGCT TGAACGATAA  | 2760 |
|    | GCAATGATTG TTCCAGTATC CGGCATGAAA TCATTTAACG GATCTTCTGT TGTAAATACGA | 2820 |
| 10 | CATTGGATGG CATAGCTTAA TGTGTATAATA TCTTTTGTGT CGGCATATAT AATCTCTTCA | 2880 |
|    | CCAAATAAAT CGGCACCTGC TGCAACTAAA ATTGTGTCTT TAACAATATC AATTCCGTGT  | 2940 |
|    | ACCATCTCTG TAATTGTATG CTCTACTTGT ACACGAGGGT TAACITCTAT AAAGAAGAAT  | 3000 |
| 15 | TCGTCAACAG ATACTAGAAA TTCAACAGTA CCTGCATTGA CATATTTAAT ATTTCCATC   | 3060 |
|    | AATTGAATTG CAGCATCACA AATACGTTGA CGTAATGTTG ATGATAATCC AACTGATGGT  | 3120 |
|    | GCAACTTCTA CAACTTTTGT ATGACGACGT TGTACTGAAC AATCACGTTT AAATAAGTGT  | 3180 |
| 20 | ACGATATTTT CATGTTCTGC ACCTATGACT TGTACTTCAA TATGCTTTGG ATTATCAATG  | 3240 |
|    | TATCTTTCTA TGTAAACTTC ACTATTACCA AATGATTTT CAGCTTCTGA TTTTGTCTTA   | 3300 |
|    | TGGAAGCAT CTCTAATTC ACTTTCTTCA CGAACGATTC TCATACCTTT ACCGCCGCA     | 3360 |
| 25 | CCACTTGTGG CTTTAATCAT TAGCGGGAAG CCAGCTTCTT CTGCAAAATC TTTTGTAAAT  | 3420 |
|    | TCATATGATT TAATTGGACC GTCTGTACCA GGAATAACTG GTAAATCTGC CTGATAGCC   | 3480 |
| 30 | GTGTACGAG CTTTAACTTT ATCTCCAAC ATATCTAAAT GTTCTAAATG AGGACCAATA    | 3540 |
|    | AATTAAATTC CTCTCTCTGC ACAACGACGC GCAAAATTGTT CATTTTCACT TAAAAATCCA | 3600 |
|    | TAGCCAGGAT GAATCGCATC CACATTCGCT TGTTTTGCTA CATCAATGAT ACGCTCAATA  | 3660 |
| 35 | TTTAAATAAC TTTCAGCAGG ACCTAAATCA CTTCCAACTA AATAGGATTC ATCTGCTTTA  | 3720 |
|    | TATCTATGTA ATGAACCTTT GTCTTCATTC GAATAAATG CAACTGTGCT GATGTCTAAT   | 3780 |
|    | CTGTGCCGCC CTCTGAATAT ACGAATTGCA ATTTCTCCAC GGTAGCAAC AAGTAACCTT   | 3840 |
| 40 | TTTATTGTGT TCAATAGCGA TACACTCTCT AACTATTAG AATTTTCTAA CTAATTAGAT   | 3900 |
|    | AATAAAATTT TATCTTAAAG CGCTCTGTGT TGCTATAGT mTGTTCmAA TTTTCAAAAT    | 3960 |
|    | TTAAATCTCT tGAGACAATT AAAaCCyCCG CTTmGAAAT AATAATTCTA AAAATGACTA   | 4020 |
| 45 | TGCAACAACA GGTAGTTCCA CGTTTTTGTG GTGAACATTT TTCGATTTCT ACAACTCTAA  | 4080 |
|    | AAAATTAAAA ATAAAAATGC AAAACATCAA CATTTATTAT CAATAGCGAT AACTTTATCT  | 4140 |
| 50 | TATCATCATG ATTCTAATTT CGCCACCACA TTTAGTAATT TTTAGTCATA AAATTAGTT   | 4200 |
|    | ATAATTATAC GTTGTTTTGT TTATAAAATT TGATAATCak GAGTAATCtC GTAATATCAA  | 4260 |
|    | AACaAAAAGG AAGTTAAGCG TTGTTTGGTT GcCTAACTTC CGTTATTGAA CTCATCagT   | 4320 |

TCTCGTACTA AATATTGGCT AGTATTTTT TAATTAAAT GTCTTCTTAT ATCAACTTTT 4440  
 TGTGTGTGTT TCITTCGTTG CTGGTCTACT TIGATTTGTT TACCTACAAT CAGAAGTAAA 4500  
 5 CCCATAGCAA TACTTAAACT AATCATTGAT GATCCACCAA AGCTGATAAA TGGCAATGGC 4560  
 ACACCAGTTA ATGGAATTGT TGCCGAAATA CCGCCAAATG TTACAAACGT TTGACTTCCA 4620  
 AAGTATGTGG CAATCCCAAC ACACACAAGT TTATAAAAA ATGATGATGT TTTATTTGCA 4680  
 10 AACTGGAAGG CACGATATAC AATAAAGAAC TCTAAAGTAA TAAC TAGCAA TCCTCGSATT 4740  
 AAACCTAATT CTTGCGAAAT AATIGCAAAA ATAAAACTCTG TATGTGGTTC TGGTAAATAG 4800  
 15 CCCAATTICA TTGCACTATT TCCTAATCCT TTTCGAAATA CGCCACCGTT ACCTATCGCA 4860  
 AGCAATGAAT TGGAAATATG GTATCCAGTT CCTGATTGCA ATTGGAATGG ATCTGTTAGC 4920  
 GTACTAAATC TGGCAGTTAA ATAACCTGGT AACCAACCAG CCATTAAATG AATGACAAAT 4980  
 20 ACTACTAAGA ATCCTAGCAC TGCTGGTATA CCAAACTCTA GGACTTTGTT 5030

## (2) INFORMATION FOR SEQ ID NO: 325:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1389 base pairs  
 25 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 325:

CTTGTTAATC CGAAAAATTAT TAGTCAATCA AATGAAACAA TAACAGACTT nGAAGGTICA 60  
 35 ATTACATTGC CAGATGTTTA CGGCGAAGTG ACAAGAAGTA AAATGATAGT TGTCGAAAGT 120  
 TATGACGTCA ATGGGAACAA AGTTGaACTA ACTGCACATG aAGATGTAGC AAGAATGATT 180  
 TTGCATATTA TAGATCAAAAT GAACGGTaTC CCTTTTACAG AACGTGCGGA CCGTATTTTA 240  
 40 ACAGATAAAG AAGTGGAGGC ATATTTTATA AATGACTAAA ATAATATTTA TGGGTACACC 300  
 AGACTTTTCA ACAACTGTTT TAGAAATGCT TATTCAGAA CATGATGTCA TTGCACTGCT 360  
 AACGCAACCA GATCGACCTG TTGGACGTAA ACGTGTATTG ACACCACCAC CAGTTAAAAA 420  
 45 AGTTGCAATG AAATATGATT TACCTGTATA TCAACCTGAA AAATTAAGTG GATCAGAAGA 480  
 ATTAGAACAA TTGCTTCAAT TAGATGTAGA TTTAATTGTA ACTGCTGCTT TTGACAAAT 540  
 50 ATTACCTGAA TCATTGTTGG CATTACCAAA TCTTGGGGCA ATTAATGTAC ATGCATCAT 600  
 GTTACCGAAG TATAGAGGTG GTGCACCAAT TCATCAGGCA ATTATCGATG GTGAACAAGA 660  
 AACC GGCGATA ACAATTATGT ATATGGTTAA AAAATTAGAT GCGGGTAATA TTATTTTCGA 720

55

EP 0 786 519 A2

ATTAGGGGCA GATTTATIAA AAGAAACTT ACCATCTATT ATAGAGGGCA CAAATGAAAG 840  
CGTACCTCAA GATGATACGC AAGCAACATT TGCCTCCAAT ATTGACGCG AAGATGAGCG 900  
5 AATTAGCTGG AATAAACCCAG GAAGACAAGT GTTTAATCAA ATTGCTGGAT TATCACCATG 960  
GCCAGTTGCT TATACAACTA TGGATGACAC TAACTTGAA ATATACGATG CTGAACCTCGT 1020  
TGAGACTAAT AAGATAAACG AGCCTGGAAC CATTATAGAA ACGACTAAAA AAGCCATTAT 1080  
10 TGTGTGTACA AATGATAATG AAGCTGTTCG AATTAAAGAT ATGCAATTAG CTGGGAAAAA 1140  
GAGAATGTTA GCTGCCAATT ATTAAAGTGG TGGCGAAAAAC ACACTAGTAG GGAAGAAACT 1200  
TATATGATAG AAAACGTGAG AAGTCTTGCT TTGACACGA TTCAAGATAT ATTAATGAA 1260  
15 GGTGCGTATA GTAACCTGCG TATCAATGAA GTGTGTGCG AAAATGAATT AAATGCAATG 1320  
GATAAGGCTT TATTTACAGA AATTGTCTAC GGAACCGTTA AAAGAAAATA TACGTTAGAT 1380  
20 TTTTATTTA 1389

(2) INFORMATION FOR SEQ ID NO: 326:

(i) SEQUENCE CHARACTERISTICS:  
25 (A) LENGTH: 2746 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 326:

TTTGCTAATA ACAATAAAGC CAAAGCCGAT TCACACTCTA AACAGCTAGA AATTAATGTT 60  
35 AAGAGTGACA AAGTACCTCA AAAAGTAAAA GATCTAGCAC AACACAATT TCGTGGITAT 120  
GCTAAGCAT TAGATAAACA AAGTAATGCA AAAACTGGTA AATATGAACT TGGCGAACTT 180  
TTAAAAATTA TAAATTTAAT GGTGAAGAAG ATAATAGTTA CTATTATCCA GTTATAAAAG 240  
40 ACGGTAAAAAT TGTTTATACT TTAACACTTA GTCCTAAAAA TAAAGATGAT TTAACCAAAT 300  
CAAAAGAAGA CATGAATTAC AGTGTTAAAA TTTCAAACTT CATCGCTAAA GATTTAGACC 360  
AAATTAAAGA TAAaATTICA AATATCACTG TTCTTACTGA TGaAAAAAGG KTTTATTTTG 420  
45 aAGAAGATGG cmaAGTTAGA TTAGTAAAAAG CTACGCCTCT ACCTGGTAAT GTAAAAAGaA 480  
AAGAAAGTGC TAAACACGTT TCAGCAAAAT TGaACAAGA GTTaAAAAAT ACAGTAACAC 540  
CTACTAAAGT TGAAGAAAC GaAGCGATrC AAGAAGATCA AGTTCAATAT GAAATACAT 600  
50 TAAAAAACTT CAAaATTwGA GaACAACAAT TCGATAACTC ATGGTGTGCA GGATTCAGTA 660  
TGGCAGCATT ATTAATGCa ACTAAAAATa CAGACACTTA TAATGCACAT GATATTATGC 720

55

|    |  |      |
|----|--|------|
|    | AAATGATTGA ATACGGTAAA TCACAAGGCA GAGATATTCA TTATCAAGAA GCGGTACCAT  | 840  |
|    | CATATGAACA AGTTGATCAA CTTACAAAAG ATAATGTAGG AATTATGATC CTTGCACAAA  | 900  |
| 5  | GTGTATCTCA AAACCTAAT GACCCACATT TAGGACATGC GCTAGCAGTT GTTGGTAATG   | 960  |
|    | CTAAAATTAA TGACCAAGAA AAACCTATTT ACTGGAATCC TTGGGATACA GAAATTATCAA | 1020 |
| 10 | TCCAAGATGC AGATTCAAGC CTATTACATT TATCATTCAA TCGTGATTAT AACTGGTATG  | 1080 |
|    | GTTCAATGAT AGGTTACKAA AAAGTAATAT AGATATTGAT TAAAGGCAGG TAAAACTATG  | 1140 |
|    | TATCAACTAC AATTATATAA TTTAGTTTAC GACACAACCA AACTCACACA TCTAGAACAA  | 1200 |
| 15 | ACCAATATCA ATTTATTCAT TGGTAATTGG AGTAATCATC AATTACAAA ATCAATTGTG   | 1260 |
|    | ATACGTCATG GCGATGATAC AAGTCACAAT CAATATCATA TTCTTTTAT AGATACGGCA   | 1320 |
|    | CATCAACGCA TTAATTTTTC ATCTATTGAT AATGAAGAAA TCATTATATAT TCTTGATTAT | 1380 |
| 20 | GATGATACAC AGCATATCCT CATGCAAAAC TCATCCAAAC AAGGTATTGG CACTTCGCGA  | 1440 |
|    | CCAATCGTTT ATGAGCGCTT AGTATAACTA ATTTAAATGA TTTCATTCA TAAAGCGGGT   | 1500 |
|    | TGGCGAGAAT TCAATTTCTC ACCAGCTCGT TTTTTCATTG TAATAATAAT CTTTAAACATT | 1560 |
| 25 | TATTCCTTCT CTATTAAATT TTCTCAAAC ATCTTATCTT TATGATAATT AATTAAAAATG  | 1620 |
|    | CCCTTTTAAA TTTCTTATAA ATAAAAAGC CACCTATCGT CGCTAATAAA CGACGCAAGT   | 1680 |
| 30 | GACTTAATAT CATATTCAAA ATAACCTATG GGAATTTAGG GAATTGATCG AAGTCAGGAT  | 1740 |
|    | CACGTTTTC TTTAAACGCA TCACGGCCTT CTTTCGCTTC ATCAGTTGTG TAATAAAGCA   | 1800 |
|    | ATGTTGCATC CCCAGCCATT TGTGTGTAAC CAGCTAAACC ATCTGTGTCA GCATTCTATG  | 1860 |
| 35 | CTGCTTTAAG GAATCGTAAC GCTGTTGGTG AGTGTTTCAT AATCTCTTTA CACCATTGCA  | 1920 |
|    | CAGTTTCATC TTCAACTTTC TCTAAAGGTA CCACTGTATT TACTAGACCC ATATCTAAAG  | 1980 |
|    | CTTCTTGTGC ATTGTATTGA CGACATAAGT ACCAAATTC ACGTGCTTTC TTATGTCCAA   | 2040 |
| 40 | CGATACGTGC TAAATATCCT GAACCATAAC CCGCATCAA TGAACCTACT TTAGGACCAG   | 2100 |
|    | TTTGTCCAAA AATAGCATT A TCAGCAGCAA TCGTTAAGTC ACAACACA TTTAGTACAT   | 2160 |
| 45 | TACCGGCACC TACAGCATAA CCTTTTACCA TCGCGATAAC CGGTTTGGGA ATAATACGAA  | 2220 |
|    | TTAAACGCTG TAAATCTAAT ACATTTAAGC GAGGGATTG GTCTTCACCT ACATAACCAC   | 2280 |
|    | CATGTCCACG TTTCTTCTG TCACCACCAG AACAGAATGC TAAATCACCT TCACCAGTTA   | 2340 |
| 50 | ATACGATAAC TGAACGTTT TGATCATCAC GTGCACGTGA AAATGCGTCA ATCATTTACG   | 2400 |
|    | CAACTGTTT AGGTGTAAAC GCATTGCGTA CTTACGGGCG ATTTATTGTT ACCTTAGCAA   | 2460 |
| 55 | TCCCTTCGTA AAATTCATAT TTGATTTTAT CATATTCTCT AAGTGTTTCC CATTGTCTGT  | 2520 |

TCCACATGAA TTGTATGACC TGTAGCAGAA ATTAATTTAC ATTACTATT AGGAATTAAG 2640  
 TTGCCAATTT TTTTCGCAAT CTGTACAAAT TTTTCATCAT ATTCTCCAGC TAATATTAAAT 2700  
 5 GTTGTACTTT TAATTTCTTT CAGCGCGGCC ATAAGTTTGG CATTTG 2746

(2) INFORMATION FOR SEQ ID NO: 327:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 900 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 327:

TAATGTTTAG TTTATTAACA GTAAGTTCGT ATATCAATGT TTAGTGCTCC CAAAAATTGA 60  
 20 AGTTTGAATT TTAAGAAGCAT CTGTGAGAAT TTAGTTGTAT TTTTTTCAA GAAATTCATT 120  
 TTGATTATTT TTGATAATGA GCATTTTAAT AGTAATACAT GTTTATAGTG TGTAGTATAT 180  
 GTCTATACTA GTAGTAACTA TATAGAGAAA GTAGGAATAA ACTATGTCAC AAGATGTAAG 240  
 25 TGAATTAAAGT AAGCAACCAA CGCCAGATAA AGCAGAAGAT AACGCATTTT TCCCATCACC 300  
 ATATTCCCTT AGTCAATATA CAGCACCTAA AACAGATTTT GATGTTGTTG AACACAAAGG 360  
 30 TGCCCTATAA GATGGTAAAT GGAAGATATT GATGATTGCT GCTGAAGAGC GATATGTATT 420  
 ATTGGAAAAA GGAAAAATGT TCTCTACGGG TAATCATCCT GTTGAAATGT TATTACCTTT 480  
 ACATCATTTA ATGGAAGCAG GTTTTGACGT TGATGTTGCG ACATTATCTG GTTATCCAGT 540  
 35 TAAATTAGAA TTATGGGCTA TGCCAACTGA AGACGAGGCA GTTATAAGTA CTTATAATAA 600  
 ATTGAAAGAA AAATTTAAAC AGCCAAAAAA ATTAGCAGAT GTGATTAAAA ATGAATTAGG 660  
 ACCTGATTCA GACTATTTAT CTGTCTTTAT CCCAGGCGGA CATGCTGCAG TTGTTGGTAT 720  
 40 TTCTGAAAGT GAGGACGTTT AACAAACATT AGATTGGCCA TTAGACAATG ACCGCTTTAT 780  
 AGTTACATTA TGTCTGGGAC CAGCAGCACT ACTTTCAGCA GGCCTTAACA GAGAAAAATC 840  
 TCCATTAGAA GGATACTCTG TTTGTGCTCT CCCTGACTCA TTAGATGAAG GTGCAAAATAT 900

(2) INFORMATION FOR SEQ ID NO: 328:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3642 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

|    |             |            |            |             |             |             |      |
|----|-------------|------------|------------|-------------|-------------|-------------|------|
|    | ATCTGAhTnG  | AGATAGTGAT | AATGTGTCAC | CcAATITTAa  | AcCtITATTT  | TTAAGcGtTT  | 60   |
|    | CaTCaGCTAA  | CaCtTCaTTa | TcTtTAGtCG | tTtTATGcCC  | tTcTATTAaA  | CtCGGaAcTtA | 120  |
| 5  | AAATGATGa   | CtTTTCAACa | CCGAACaCTA | AAACaTTGtC  | tTTTtTGATGG | CcATTAGACA  | 180  |
|    | CAATTTCCCC  | TGTTTGCTTC | AAAGTAGCTT | GCTTCITGTA  | TTTATTTTCA  | ATATCTTTCT  | 240  |
|    | TGTTAAAAAC  | AGATTGTTCG | ACAGTTTGAT | TGGCATCTTT  | ATTTAGAACA  | ATGGCATCTG  | 300  |
| 10 | CTTGCCACTT  | ATCAATGCCT | TCTTTATTCA | TATTGATAAG  | ACCAITTCGCC | AATCCAGATA  | 360  |
|    | ATAAAAATAG  | CAAGTAACTA | ATCATCGTTA | ACACACCAAT  | AATTAGTCCA  | AACTTCAATT  | 420  |
| 15 | TGTTGCGCGC  | TATTTCAATC | CAAGCTAAAA | ACATGCATTT  | CTCTCCCTAC  | TACTATGATT  | 480  |
|    | TAAACATTGT  | TTATATTCTT | AGATGCACGT | ACGTCGTGTT  | GCGCTCTGTA  | ATGTTATACA  | 540  |
|    | TACACTTATC  | CTTCATTATA | CCGAACTTTT | TTATATTAAA  | AGCAAAATTA  | TGGAAAATGC  | 600  |
| 20 | AAITAAITGT  | CTATTATTTT | TGTACGGTAC | ATTTAAAAAT  | AAGGATCAAT  | TTAAAAACGC  | 660  |
|    | CTACATATAC  | CTTTAAGTAC | ATGTAGACGT | CCAATTCATA  | TATTATTTAA  | CTTCGCTGT   | 720  |
|    | TTTAGGATCG  | AATTCCTTAA | TAGCATTTT  | ACGTAATTTA  | TCTTTTGCTT  | TtTCACTTGC  | 780  |
| 25 | TTTATAGTTA  | TGTTGTAA   | TGCTAGCTTC | CCAACCTACCA | TACATTTGGT  | TAGGGAAAAT  | 840  |
|    | GATATATTTT  | TTACCGAAAT | CGTCTTTATG | TTTTTCAATT  | AATGCTTCAC  | GAGATTGAGC  | 900  |
| 30 | TGTAGCTTCT  | TTGGATCTG  | TAAAGTCTAA | TAAATTATCT  | CCAATATAGCA | TGACAAGTTT  | 960  |
|    | ATGATCCTTT  | TGAACCAATT | GTCTGCGTGA | TTCTTTACTC  | TTATCATCTT  | TACCTTTTAG  | 1020 |
|    | TAAATATAGA  | CTCTCTTAG  | CTTGAGGGAT | ACCTTGTTGT  | TTTAAAGTTCT | TTTGTTGTTG  | 1080 |
| 35 | CTTTAAATCT  | TTTTCTTTAT | CTCTATCAGA | AATATAGTAG  | ATATCGACAC  | CTTTTTTGTC  | 1140 |
|    | AGCATATTTT  | AAGAATTCCT | TTGCGCCATA | GACAGTTTTA  | GCTTTAGCAG  | CTTTGATCCCA | 1200 |
|    | TTCATGCCAA  | CCTTCTGGGA | AAGGTTTATT | ATGTATTGAT  | GCATAGCCTT  | GATATGGAGA  | 1260 |
| 40 | ATTATCTAAA  | ACTGTTTCAT | CTAAATCCAA | AGCAATAGCT  | AACCTATGTT  | TACCTTTATT  | 1320 |
|    | CTTTTTTAATC | TCTTTATCTA | ACTGTGCTT  | TGCACTGTGA  | TAACCTTGTA  | AATATAATGC  | 1380 |
|    | TTTTGCTTCA  | GCTGAATTTT | GATACCAAGC | CACTGCCATA  | ATATTTTGAT  | TACCAAGATT  | 1440 |
| 45 | CGCCTTTTGT  | GATGCTGGTA | TAGAAGCTTG | TTGCGTTTGT  | TGAACCTCAG  | CAGAACTTTT  | 1500 |
|    | GGCAAAACGT  | GTAGAATTTG | TCGTTTGTTG | TGCTGAAACT  | GTAACCGCTA  | CCGATAATGA  | 1560 |
| 50 | TGCTATTGCA  | ATATACTTTG | AAATTTTATT | CATCTTATCA  | CCTCATGATT  | AATATTTAAA  | 1620 |
|    | ATACAGTTAA  | AATTATAAAT | GCATTTATTT | AATATTGCTA  | TACTATGAAA  | AGATATTTAA  | 1680 |
| 55 | TATTATTCTT  | TGgAAAAGCT | AACAAATATG | TGAACATTTT  | ATATAAGCAT  | GATTTAATGG  | 1740 |

EP 0 786 519 A2

|    |             |             |            |             |             |             |      |
|----|-------------|-------------|------------|-------------|-------------|-------------|------|
|    | GTGACTAGGC  | CTTCCTATCA  | GACATATTCA | CTCATCCACG  | TATCATTATG  | TGTACAGTGT  | 1860 |
|    | GCTATCTCTT  | ATTTTACCTAT | TGGAACAACC | ATAAACTCAT  | CCATAGTTTA  | CCTTTTATAA  | 1920 |
| 5  | ATAGCAGTCC  | TCACATCATAC | AATTTCTCAT | AAAAATCACA  | ACGCTCCAAC  | GTATTTCCAA  | 1980 |
|    | CTTACTTTCA  | CCTATTTTAA  | TTCATAAAAA | CGACACTTTA  | ATTGTCATTA  | TCCAATAATA  | 2040 |
|    | GCAAGACGTT  | ATTATTGCAA  | TCTTTTTTAT | AAAATAATAG  | AATCATAGTA  | TTGTCATTTA  | 2100 |
| 10 | AAGATAAAGT  | AAGAACGTTT  | TTATTTTTCa | CAITTTTTTAA | ATTATTATGA  | ATATCTAGTT  | 2160 |
|    | TTAGGAAGGA  | AATTACATTG  | AAAAACAAG  | TTATTATTTT  | GGGCTCATG   | TTATTTTCAC  | 2220 |
|    | TATTTTTTGG  | AGCGGAAAT   | TTAATATTCC | CGCCCATGCT  | TGGCCATACA  | GCGGGTCAAA  | 2280 |
| 15 | ATATGTGGAT  | TGGTATGCTA  | GGCTTTGCCC | TTACAGGCAT  | AThACTCCCC  | TTTATTACTG  | 2340 |
|    | TTATTTGTGT  | TGCATTTTAT  | GATGAAGGTG | TTGAAAGTGT  | AGGCAATCGT  | ATACATCCAT  | 2400 |
| 20 | GGTTCGGGTT  | TATTTTTGCT  | GTCGTGATT  | ACATGCTCTAT | CGGAGCATT   | TACGGTATTC  | 2460 |
|    | CACGTGCTCG  | AAATGTCGCG  | TACGAAATTG | GTACAAGACA  | CAITTTTACCT | GTGCATAACC  | 2520 |
|    | AATGACATT   | AATTATATTC  | GCAGCAATCT | TTTTTGCCAT  | CGTTTACTCG  | ATTAGTTTAA  | 2580 |
| 25 | ATCCATCGAA  | AATCGTTGAT  | AATTTAGGTA | AATTATTAAC  | ACCGTTATTA  | CTATTAAATG  | 2640 |
|    | TGCTCTTAT   | AGTATTGCT   | GTCATTTTCA | ACCTGAATC   | TGCACTAAGT  | GCACCTAAGG  | 2700 |
|    | ATAAATATAT  | AACACATCCT  | TTCAITTCAG | GAAGTTTGA   | AGGCTATTTT  | ACAATGGATC  | 2760 |
| 30 | TTGTGCTCG   | GTTAGCTTTT  | TCCGTAGTCA | TTGTCAATGG  | CTATAAGTTT  | AAAGGCCTCA  | 2820 |
|    | CAGATCGCAT  | GAAAATTTTA  | AAATATGTCT | GCITTTTCAGG | TCTTATTGCA  | GCCATATTAC  | 2880 |
| 35 | TTGGAATGAT  | TTACTTTGCA  | CTTGCATACG | TTGGGGCATC  | AACAGCTCCA  | GGAAGCTTTA  | 2940 |
|    | AAGATGTTAC  | AGATATATTG  | ACGTACAAC  | CATTACGATT  | ATTGGTTCG   | TTGCGTAACC  | 3000 |
|    | TCGTATTTGG  | AATGACGGTT  | ATCCTTGCA  | GCCTAACAAC  | ATGTATAGGA  | CTCGTCAATG  | 3060 |
| 40 | TTGCGGCACA  | TTTACTAAGA  | AACACGTACC | TAAGTTTTCT  | TATAAAATAT  | TGCACTTTAT  | 3120 |
|    | TTTCTCTATC  | ATAGGGTTCT  | TATTTACAAC | ACTTGGTTTA  | GAAATGATT   | TAAAAATTGC  | 3180 |
|    | TGTCCTCATTA | TTGACTTTAA  | TATATCCCGT | GTGATTGCA   | CTTGACTACA  | TATCATTTGC  | 3240 |
| 45 | TAACATGTTT  | AGCACAATTCA | GATTCAGTTG | GGCCTATCGA  | CTCGCAACTG  | TTATTACATT  | 3300 |
|    | GATTATTTCA  | ATTTTACAAA  | TACTAAATAG | TTTCAACTTA  | TTACACGGTG  | TTATTTTGAA  | 3360 |
| 50 | ATCGTTTATG  | ATGTTACCTT  | TAGCAGATAT | CGATTTAGCT  | TGGCTGTAC   | CATTCTAGCT  | 3420 |
|    | CTTTGCTATT  | ATCGSTTTCA  | TAATCGATGT | ATTTATACGC  | CGTCCGAAAC  | AAGCGACAAC  | 3480 |
| 55 | TTAATAAATG  | CTCACTGCCT  | AGTAATGATT | GACCCATCGT  | TACTAGGCTT  | TTtttATATGA | 3540 |

# EP 0 786 519 A2

TAAAAATCC TAGCTGTIAT TCAAAAAATAT TAGTTTTTAA AA

3642

(2) INFORMATION FOR SEQ ID NO: 329:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 2187 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 329:

|    |   |      |
|----|---|------|
| 15 | TTATTGATAT TGAAAAATCA AAAACTGCAA CACATCTTAT CACTAAAAAC CAACACTATC   | 60   |
|    | ATCAAACCTGA CACTCATTTT GAACAGTATA AGAAATTTAT TTTAGATTCA GGTATATCAT  | 120  |
|    | CAACACAATT TGTATATAAT AACCTGCTG TAAGCGGATT TAAATATACT AATGATGGTA    | 180  |
| 20 | AGAATCCAAT TCAATTATCT GACATAGTGT ATCACTTAAT CGCATTATTA CGATATGGCG   | 240  |
|    | GTGGCATTAG CTATCAATTA TTAGATGACC ATTCAAATTA TATTTCCTTG TACAACAAAT   | 300  |
|    | ATGGTAGCCC CCTACCATTAT ATGCATTIAT ATAAAAATGTT TAGACCTTTT GTTAATGAAG | 360  |
| 25 | ATATTGAAAT TACAAATAAT TATGTATTGA GTCGTAAGA TAATAACTAC CATTTCTTAT    | 420  |
|    | TATTCATAAA AATTAAATGAT CGATATATGT CAGACGTAAA ACAAGATTTC ATTTTCATA   | 480  |
| 30 | ATGAATTACC TCAAGACTCT TTGATGATTA TTAACAATT GAATCATGAA CATGGTTCAA    | 540  |
|    | TTCAACATTT GCTTCCAATA AGCGATCAAC TTGTTTATAT AGAAAAAGAA ATTTTAGATG   | 600  |
|    | AATTAGACAA AACGAATTAC CCTAAAACGG AGCTTGCAGT TCAAGAAGAA ACTGGTCGTA   | 660  |
| 35 | CATTTGAAGT CAAGTTAAAT CACGACGAGG TTAATATATAT TTGCTTTAAA CCAAGCTAAA  | 720  |
|    | TACTAACAGT CCTCTTGTTG TTAGTTTCTT ACGTTAAAGG CTATTTATAT CATAAGGAGA   | 780  |
|    | TGATTTGTAA TGAGTAACCT ACAAGCAATT CAAGCAATTG AAAACGTGTT AGTAACGTCA   | 840  |
| 40 | AAAGTTGGTG TATTATCAAC TGCATATAAT AATAAaCCTA ATAGTAGATA TATGTCCTTT   | 900  |
|    | TATAATGATG GTCTTACTTT ATATACTAAA ACGAATATCC ATTCGCTAA GGTCAAAGAA    | 960  |
| 45 | ATTAAGATA ATCCAGCAGC ATATGTTTTG TTAGGCTATA ACGACACAAC TAATCGTAGT    | 1020 |
|    | TCGTTTGAGA TGGAAAGCGAC AATCGAAATC GTTACAGAAC AAGAAGTGAT TGATTGGCTA  | 1080 |
|    | TGGAAACAC AAGACAAAAG CTTTTTCAGT TCAAAAGAAG ACCCAGAGTT AGTGTTTTTA    | 1140 |
| 50 | AGAGTAGTTC CGCAATCCAT TAAGCTAATG AATGATAAAT CATTAGATAC ACCTATCAAA   | 1200 |
|    | ATCGATTTAT AACACAAAGT GTATATAGGA AATAACTTTT ATGAATTCTA GATATAACAA   | 1260 |
| 55 | TGTTAAATAC TTAATAAATC TCGCTATAAT TAAAGTGTTT AATATGTTTA CAATTCAATT   | 1320 |



TTTAAATTA TTTTATGTA TATAAACT GCATTGCAA ACTGTTGCAC TTTTAGGTAT 1440  
 AACAGAATTA ACTACATTTA AGGAGATTGA TGAACCATGA AAAAGAAAAA AGGTTTGGT 1500  
 5 CTGTGATTA GTTTAATCGC CATCATGTTA ATTGTATGTA TTGTATTAGT AATCATGATG 1560  
 ATGACTGGCG GAAAGAAAGA TACATACTAT GGAATTATGA AAGATAATAC TACTATTGAA 1620  
 10 AAAATGATTA GTGAAAAAGA TGAAGTATT GAAAAAATG TTAATTTACC TTCAGATTCA 1680  
 GATGTTAAAG TTAATAAAGG TGATTTTCTA ATTGTTTATA AATTAGCAGA TTCAGATAAA 1740  
 ATTGTTAAAG TTAATAAAGT TGACCATGAC GATGTACCAC ATGGTTTAAT GATGAAATTT 1800  
 15 CATGACATGG GCAAAATGCA CATGAAACAC TAATTGTAAT TTAATTACA AATTTTAGTT 1860  
 GCCATCAAGG TATATACGAG TAAAGCAGC GGTAAGTTGA TTTCCAATT GGAATCATTT 1920  
 TACTGCTGCT TTTTATATTT GAAATCTTT CATATTGAAT AGCTCCACTT GCCGTTGCGC 1980  
 20 TgcgCTTTGC GCATGCATAA AAGCCCCCTA CAACCTGAGG TCACTGCGCT CCGGTTGCGC 2040  
 TgcgCTTTAG CGCATGCATA AAGCCCCCTA ACAACCTGAG GTCACTACGC TTCGGTTGCG 2100  
 CTGCGCTTTA GCGCATGCAT AAAAGTCCCT AACACCTGA GGTCACTAGC CTTCGGTTG 2160  
 25 CTTGCGCTTT AnCATGGCCA TAAAGC 2187

(2) INFORMATION FOR SEQ ID NO: 330:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1788 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 330:

CChCCTTTTT AAACCTGGnG AAATGmCAG tTTTGATGGt ATTGGGtCTT TTAGTATTAT 60  
 40 GCTTAAGATA GAGTGTAATA CGCTCCTGTT GTTCTTTATA GTATATTGCT TTTGTTTTT 120  
 CTTTTTCGTC ATATTTCACT TTTAAATAAA TGACTGATGC AACTATACAT ATACATAAAA 180  
 45 TGACACCAAT AATTATAAAA ACATGTTTTT TGTTCAATTA AGTAACTCCT AAAATGTGGT 240  
 GGAATAAGAA AAAATTTTTA TAATCTATAA TTATGAACAT TACAGATTAT AAACCATAAC 300  
 ACTAACATCG TCGCCTTCAT TAACTTGATT GTTAAATCA GCAATTACTG AGAATTGTAC 360  
 50 AAGGAATAAT TGCCATTATAT GCCCTCATGT AATTATTGCC TTACTAACAT TAACAAAATC 420  
 GATAGCTATT ACATTAAATG CCTATACCCC AGACCTCAA CACCTTTTTA TACAGGACGC 480  
 55 ACTGTAATTT CATyyACGTT AACGTGCTTT GGTGTGTGTA ATGCATATAA TACTGCTTCT 540

GTATCTACCA TTCCTGGAGA AATGCTTGTT ACTTTAACGC CTGTCCTTGC CAACTCTTTT 660  
 TCTAATCCTT GAGTAATAGT GTGAACCTGCT GCTTTCGTCG CACTATAAAT CGTACTACTT 720  
 5 TTCGTTACTT CAAAGCCAGA AATAGATGcA ATGTTAATAA GATGGCCACT TGATTGTTCT 780  
 AACATAGTTG GTAATGCAGC CTGTGCCGTA TATAAAGTGC CTTTGATATT CACATCAATC 840  
 10 ATACTATCCC ACTCATCTAC TTGATAATCA GTAATCTTAG ACGACAACAT TTGCCCCGCA 900  
 CTATTGATAA CAATATCCAA ACCACCGAAT GTTTGTTGTG CAATTTTATT CAATTCATCG 960  
 ACTTCTCTT TATTGCTTAC ATCTGTTGGC ACTACCTTCA CACTATCTTG TGACAATTGA 1020  
 15 TTCGCTAGST TTTGTAATTT TTCTTTATTT CTACCTGCTA AGACAACCTT TGCCCCCTTCT 1080  
 TCATGTAGTA ATGTTGCAAT TGCTTCTCCA ATACCACTAC CTGCACCTGT AACTACTGCT 1140  
 ACTTTATCTG TTAATACTGT CATAATGATC GACTCCCTTG ATTCTTTTTA TTTTTCAGG 1200  
 20 GTAAATCATA AATACATATT ACTTTTAAAA AGCGTATCAC AATTCATATA ACGGTCATAA 1260  
 TAACCTCGCTT CATTTTCATA GATAAAATAC ATTACAAGCC ATTCGAAACA TACAATTAAT 1320  
 CGTTGCTTAT ATTTTITATT TTTAAAAATG TTGAAAAATC GTCACCTCTT TATTGTAAAA 1380  
 25 ACATTATATT AGTAATAAAG TTAATACTGT GtATTtATCA TTCGATTGAA TGATTAGAGG 1440  
 GAGGAATAAA AcgTGACATA TCATGAGCGT GTTTTAGCAT TAAGAGCAGA AAGTAAAAGA 1500  
 30 ACCGCATTG ATTTTCGATT CGAAGATTTA TTTAGCAAAG AAGAATGGcT AAGTATGTCT 1560  
 CTTGCAGAAA GACAAAAAGC TGAAAAAGCA TTTGCACACG AgTTAAAAAT ATGGACGATG 1620  
 TAAGAATGCC CTTCCTCAAGT GTCATGACG CCCAAGTAAA ATTATATAAT GTTGATATATT 1680  
 35 CTTATAACGG CaTTAAACGT AATTTTAAAC AAGTTGAAAA TGGAAGGATT CTAATATCAT 1740  
 TTCGTTTATA TATnGCAGAC CATGGATAGA ATTTInTATG GtnAACTCC 1788

## (2) INFORMATION FOR SEQ ID NO: 331:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1341 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 331:

TAAGCCAAAT CCATTGTTAC CAGAAGAAGT TCGCTTCATC TTAAACAAAA TGGGCTCTTAA 60  
 AGAAAAAGACT ATCGATGTTG AACTCGAAGT TGGCGAGCAA GTTCGTATTA AATCAGGTCC 120  
 ATTTGCGAAT CAAGTTGGTG AAGTTCAAGA AATTGAAACA GaTAAGTTTA AGCTAACAGT 180

EP 0 786 519 A2

|    |  |      |
|----|--|------|
|    | GCTTTAATTA ACAATTAAG TTATTAACT AACAAAAGA TAAAAAGAG TATTGATTTT      | 300  |
|    | TTAATTAGAA AAGTGTTAAA ATTATGTGGT cGcGCTTTTA GAGCGCCCAT TTCGTcACGA  | 360  |
| 5  | AATGTTAAGA GTGGGAGGGC AAAACTGAGC CCTGTGACCA CATCAGGATA TCAAGGAGGT  | 420  |
|    | GCACATCGTG GCTAAAAAG TAGATAAAGT TGTAAAATTA CAATTCCTG CAGGTAAAGC    | 480  |
| 10 | GAATCCAGCA CCACAGGTG GTCCAGCATT AGGTCAAGCA GGTGTGAACA TCATGGGATT   | 540  |
|    | CTGTAAAGAG TTCAATGCAC GTAACAAGA TCAAGCAGGT TTAATTATTC CGGTAGAAAT   | 600  |
|    | CAGTGTTTAT GAAGATCGTT CATTACATT TATTACaAAA ACTCCACGG CTCCaGTATT    | 660  |
| 15 | ACTTAAAAAA GCAGCTGGTA TTGAAAAAGG TTCAGGCGAA CCAACAAAA CTAAAGTTGC   | 720  |
|    | TACAGTAAGT AAAGATCAAG TACGCGAAAT TGCTAACAGC AAAATGCAAG ACTTAAACGC  | 780  |
|    | TGCTGACGAA GAAGCAGCTA TCGTATTAT CGAAGTACT GCACGTAGTA TGGGTATCGT    | 840  |
| 20 | TGTAGATAAA TTTTACGAAT ATTAATTTTG ATTACATGAT TTAACGATG AAGCAGATAA   | 900  |
|    | CAGAGATAAT AATGATGAAT TATAAATATA ATCTGAATGA CTAGATTAAT GATTGATTTA  | 960  |
| 25 | TTCATAAGAT TAATTCTTCT GTTGCTGTCy CTAACTTGC ATATAGCAAG TAATGTGGGA   | 1020 |
|    | GGAAATCCG CTAAACAC TAAAGGAGGA ACTATAAATG GCTAAAAAG GTAAAAAGTA      | 1080 |
|    | TCAAGAAGCA GCTAGTAAAG TTGACCGTAC TCAGCACTAC AGTGTGAAG AAGCAATTAA   | 1140 |
| 30 | ATTAGCTAAA GAAACAAGCA TTGCTAACTT TGACGCTTCT GTTGAAGTTG CATTCGGTTT  | 1200 |
|    | AGGAATTGAT ACACGTAAAA ATGACCAACA AATCCGTGGT GCAGTTGTAT TACCAACCG   | 1260 |
|    | AACTGGTAAA TCACAAAGTG TATTAGTATT CGCTAAAGGT GACAAAAATTG CTGAAGCTGA | 1320 |
| 35 | AGCAGCAGGT GCTGACTATG T  | 1341 |

(2) INFORMATION FOR SEQ ID NO: 332:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 5136 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 332:

|    |  |     |
|----|--|-----|
|    | CTCTAAATCT TCAATAGGTA ACTTCATTTT AATAATCCCC ATTGAACAA TTGCTTCTTC   | 60  |
| 50 | ATCATTGACA ATTTCTAAAA CTTCACCCCTT TTGACCGTAA GAKAATACTT TkACTTCATC | 120 |
|    | ACCAGCAACA ATTTTATCGT ATTTTGTGCTT TTGTACATTI TGCTTTATAG ATTTGCTTC  | 180 |
| 55 | ATAATGATCA TCTAATCGTT TCTTCTTATC AATCAATTCa TGTTCTTTAA CATCTGCACC  | 240 |

|    |            |             |            |             |             |             |      |
|----|------------|-------------|------------|-------------|-------------|-------------|------|
|    | CTTAATCTTC | TGATTGCTT   | TkTCITTCGC | TTCTCTATT   | AGAGACTTTT  | CATAAITTTG  | 360  |
|    | GAATTGTTGA | TACTGCTTAG  | ATAAATCATC | ATGCACTTGT  | CCGCKTsKtT  | TACAAGACGG  | 420  |
| 5  | TCCAGkTCTA | ACCTCTGTGT  | CTCTACACGT | TTGTAATTAC  | GCTCTAATGA  | TTCAATCATT  | 480  |
|    | TCATTTATTT | CTTTTCATC   | AGTACCAATC | ATCGTCTTAG  | CCTTATTAAAT | AATATTCAAA  | 540  |
|    | CTAAGACCTA | ACTTTTTagA  | AATGTCAAAA | GCATTTGAAC  | GACCCGGCAC  | ACCCATTAAAT | 600  |
| 10 | AACITATAGC | TTGGACTCAA  | AGTATCTACA | TCAAATTCTA  | CACTCGCATT  | CATAACGCCT  | 660  |
|    | TCTCGATTAT | AACIATATGC  | TTTAAGTTCA | GGATAGTGGG  | TGTTTGCCAT  | TACTAGAGAA  | 720  |
| 15 | CCAATTTTTC | TAACATGATC  | TAAAATGCTC | ATTGCTAATG  | CAGCACCTTC  | ACTCGGATCT  | 780  |
|    | GTACCTGCAC | CTAATTCATC  | AAATAAAACT | AAACTATGTT  | TGCTGCGATG  | CTTTAAAAAT  | 840  |
|    | TCAACTATAT | TCGTCATATG  | AGATGAAAAA | GTTGATAATG  | ATTGTTCTAT  | TGATTGTTCA  | 900  |
| 20 | TCTCGGATAT | CGCAATATAC  | ATTTTTAAAT | ACACTTAACT  | GACTACCATC  | AAGTGTGGGA  | 960  |
|    | ATCAACAATC | CTGATTGAGC  | CATAACAATA | ATTAAACCTA  | ATGTTTTTAA  | TGTTACAGTT  | 1020 |
|    | TIACCACCTG | TATTCGGTCC  | TGTAATAATT | ACCGTTTCAA  | TATCTTCCAT  | AAATTCGATG  | 1080 |
| 25 | GTATTAGCTA | CAACAGTCTC  | ACGATTTAAT | AATGGATGGT  | ATGCTTTAGG  | TAAATATACA  | 1140 |
|    | GTACGGTCTT | CTTTAAATAT  | CGGCTTTGTT | CCTTTAATAC  | TTCTACTATA  | TCTCGCTTTT  | 1200 |
| 30 | GCGATTAAAA | AATCTAACTG  | ACCCATGACT | TGTTCTGCCA  | CAAGTAGTGC  | ATCTTTGTCC  | 1260 |
|    | GCAGCCACAT | AACCAGTTAG  | TTGCGTTAAA | ATGCGTTCCT  | TTTCAATTGC  | TTGCTCATGA  | 1320 |
|    | CGTAATCGAC | TAATTGATT   | ATTCATTICA | ACAACCTGAT  | ATGGCTCAAT  | ATACAATGTT  | 1380 |
| 35 | TGCTCTGAAG | CAGATTGATC  | ATGTACAATC | CCATTAAAAAT | CTTGCGGATA  | TTCAGCTTTG  | 1440 |
|    | ACAGGTATAA | CGTTTCTTTC  | ATTCCTAACT | GTTACAATAG  | CATCTGATAA  | TTTTTTCTGA  | 1500 |
|    | TTTGCTTGCC | TTTTAACAAAT | ACGGTCCAAA | TTTTGTCTAA  | TACGTTGATT  | CGTGCTAGAA  | 1560 |
| 40 | ATTTTACTTC | TAATCCCTTG  | CAATTCAATA | CTCGCATTAT  | CATATAAATC  | ATACGTATCG  | 1620 |
|    | CATGTTTCAT | TTATTGTGTG  | AAAAAGATCA | GTTAACACAG  | GTAATTGATT  | CATCTTGTCa  | 1680 |
|    | TCTAATATTG | GGTATTTAAT  | ACCTTCATCT | TCTTCAACCA  | ATTGATTATA  | AAATGCTTTG  | 1740 |
| 45 | AATTGATTTT | GTACTTGAAT  | TAATCTTTTT | ATCAAGTTAA  | GCTCTGATAC  | ATTTAAAAGG  | 1800 |
|    | CCGCCAATAT | CAGCCGGATG  | AATGAATGCT | GATACTTTAG  | ATAAGCCACT  | CAAGCTTGGT  | 1860 |
| 50 | AAACGATGCT | TATTATAGAT  | TTGAGCAATC | TCATCCGTTT  | CTTCATTITG  | AAAAACAACC  | 1920 |
|    | GTTTCAAAAT | TAGTAGCTGG  | CATCATTIGA | TTGACCTTTT  | CCAAGCCTAA  | GTCATAATA   | 1980 |
| 55 | GTTTCATTGG | CAACGAGTGA  | TTTTATTTTT | TCAAATTCTA  | AGACGTCTAA  | TGTTTTTTGT  | 2040 |

|    |  |      |
|----|--|------|
|    | TCACGCGATA ATGCGTTAAT CACTCTATCT TTTGTACAA ATCCTTTTTG CGCAGTTGEA   | 2160 |
| 5  | GCCTCAATTT CATAAAATCT AAATGATTGG TATGATGCGC ATCAGTGTGA ATAGTTAATT  | 2220 |
|    | TCACATTGG ATATTTACGA ACGATATCAG CGCTCAGATC CAGTCGATGT GGATTGGCAT   | 2280 |
|    | TAATTTCTAA TACTGTATTC GTTCTTCAG CTAATGCCAT TAATTGTICA ATATTGGTT    | 2340 |
| 10 | TATAACCATC TCTTCTACCT ATAATACGCC CTGTTGGATG CGCTATATGT CGCACGSTATG | 2400 |
|    | GATTGGGACA TGCATTAGCT AATCGTTCCA TAATTTGTTC TTCTGATTGG TTAAGCTTT   | 2460 |
|    | GATGAATAGC TCCAATTACA TAATCAAGTT GTGCTAAAA TTCATCATCA TAATCCAGCG   | 2520 |
| 15 | AGCCATCAGG TAATATATCC ATTTCTGTAC CTGAATAAA ATCAATTICA CTATATTCTT   | 2580 |
|    | TATCTAAAGC CTTAATTTCT TCGTTTTGTC TTAAAAGTCT TTCTACTTGT AAGCCATTAG  | 2640 |
|    | CAACACGTAA ACTTTGTGAA TGATCAGTAA TTACCATGAA TTTATAACCT TTTGCGATAT  | 2700 |
| 20 | TTGCTTCTAC CATGCTTCGA ATAGAAAACG CACCATCACT ATACGTTGTA TGCATATGAA  | 2760 |
|    | TATCACCATT AATATCATCT ATTGTAATGA TATTACTTAG ATCTTTATCA AATTGCTTAC  | 2820 |
| 25 | CATCTTCTCG CATAGCAGGT GGTATAAAAT TCACATTAAA ATGTTCATAT ATCTTGGCTT  | 2880 |
|    | CACTATCATA TTGAATTAAT GTACCATCAG CTGTTCAAT TCCATATTCA CTTACTTTTT   | 2940 |
|    | CATCAGTGCG TTTAGCAAGT TGTCGAATTC TTATATTATG TTCTTTTGAC CCAGTAAAAT  | 3000 |
| 30 | GCTGCAATGT ATGATAAAAA GCACTTGGTT CAATTAATCG AAAATCGACA CCAATCGTTT  | 3060 |
|    | CATCATCATA CGCTAATCTT AATGAAACTT TTGTGTTCCC CACTGCAACT TCTTTTACTT  | 3120 |
|    | TATTGGGAAT ATTTAATAAT TGCTGCTGCA CTGCTTTTGG GTTATCGGTA CTTAATTATGA | 3180 |
| 35 | AATCTAAATC TTTGCTCATT TCTTTAAAA GACGGAAGCT TCCTGCAGAT GAATATTGAT   | 3240 |
|    | CGATATAATT TAATGTATCT ATATAATCAA TGATTTCTTG ATTAAGTCTT CTCATTTGAT  | 3300 |
| 40 | CAATTGGATA TCTATCTTTC TTAGCACCAA GTTGTTTCAC AGCTTCTAAT ATGTTTGTGT  | 3360 |
|    | CCGTTTTCTT AGCAAAATCG CTTAATTCAC TAACTTTTCC ATTTTCACAA GCAACTTGAA  | 3420 |
|    | GTGACGCTTT ATCAACAATA TTCAACTCTT TATATAGCTT AGCAATTTTC TTGCTTCCAA  | 3480 |
| 45 | GTCTTTGAAT TTTCAAAGT GGAATAAGAC CTTCCGGAAC TTCTTCTCTG AATTGCTGTA   | 3540 |
|    | AATACGTAGA TTCACGGGTC TCACGGTAAT CATGTATTAC TTCTGCAACA CCTTTACCA   | 3600 |
|    | TGCTTTTCAA CTCGGTACA TCAGATATTT CATCTAATGG TCGTTTCATCT AATTCAAGAC  | 3660 |
| 50 | TTTGAGCTGC TTCTGATAC GCTGATATT TAAAAGTATT TTCCCTTTT AATTCCATAT     | 3720 |
|    | AAGTAGCAAT TTGTTCTAAT AGTTTGATAA CATCTTTTTT TGTCATAATA ACACCTCATA  | 3780 |
| 55 | AAAAGAAGAC CAGGACGTAT CATTAAATATA TACCTTTGTC CTGACCTCTT ATGATAATTT | 3840 |

TAGATATTTT AAGCTGATGT TGTAAATGCTT CGTTAGGATA TAATGCCAAT AGATATAACG 3960  
 TAAAGTGTA GACAATTATC GTCATAAACA CACCAACTAT CATTCGCCATT GCACGACTGA 4020  
 5 AAAATGTAAT GTTTTGATAC GCTATTATTT TATCAAAAGT TACGATAATT AGATATAAAA 4080  
 TGAACCTACA AAACAATGTA ATCATTAATAA AAGCTACAAT CGCTTCAAAT CGATTTTGTGTA 4140  
 10 GATGATTAAA ATGAAACGCA AAAAGTTGTAT TAAATGCTGT TGTTTTAGGA TAIGGAATAA 4200  
 ATACAATTAA TCTTTCTACA ATAGATTTGT AAAATTGACT GGCAATCCAC AATGATACAA 4260  
 nCGTTGCACT CAAATGTATC ATAGATAACC AAAAACCTCG TCTGAATCCA ACGATGACAA 4320  
 15 AATACACAAA GAAATGATT ATGATAAAAT CAATGACCAT TTATTGCTCA CGCTGCTGCA 4380  
 ATTTGTGAAT TGTGTGTTT AAACGTGAT TTTTTTCTT CTAGTAGTAC TTTTTCATGC 4440  
 ATAATATTCA CAGCAGTTAG TATTGCTTTT CTGGAAGTAT CTAAACCTGC TGCCTTATAC 4500  
 20 CCTAATTTCT TATTTTTATC ATCAACTAAA TGTGCTACAT ATCGTATGTG CTCTGGGTTA 4560  
 TCTTCCCCAA CAATTGTAAA AAGCTGATCA TTAATTGATA CATTACCTT GTTTTTAAAC 4620  
 TGTGTCAATT ATAATTTCTC CTGATCCTTT TTTTAAATC TAAATTCACG TTATAAAATA 4680  
 25 TGACTGGATA GTTTGTCTGA ATTTGATACT AATATTGTTA TATTGCAATT ATGATAAAAC 4740  
 AACACACAAA TCTCTATAGA TGACTTAATG TTCTTTTAT AATGAAATAA TGTAAGAAT 4800  
 30 TTTCTATTCA ATACTTTATC ATGTTTAAAT TGTGCTACTA TAACATTTTC ATAAACATTA 4860  
 TACATGACCA CTATGTATTT TGTAAGTATC CGCAATTAAAT TCTTTACAAC ATACATAAAT 4920  
 GTTTCGACG TTATTATCAT TTATGATATG ATTATTTTGT CTAAAGACAA TGAATTTTA 4980  
 35 TGAAAGGATT TACACAAATG GCGAATATCG TTTTAAATTT GTCGGATAAA GACATAACGA 5040  
 CATTAAATGC ACGCATTTC TTTGATACTG AGAATTTACC TCAAGGnATG AnAGCAGCTG 5100  
 CAAAGTATCC AAAATACAAC TGTAATATTT TACCAT 5136

(2) INFORMATION FOR SEQ ID NO: 333:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 4239 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 333:

GGCCAAAATT GCACCTCCAC TTTCTTGA ACAATCTAT TTAAACGCA CATTCGCCATT 60  
 ATGACAAAAT AAATGTGAGT nACATTGTGT TGTATTTTAA CATGACTACT AACGCAAACT 120

|    |             |             |             |             |            |             |      |
|----|-------------|-------------|-------------|-------------|------------|-------------|------|
|    | GGTATCTTTC  | AAAGATAAAA  | TCTTAATAAT  | TTCTTAGTAA  | ACTCTTTTCT | CTAGATTAT   | 240  |
|    | CACAAATATTA | TATAGACCTA  | TTTTATTTTG  | ACGTAAGTTG  | CTAGTATCTT | CAAAACAAAA  | 300  |
| 5  | ACCTTTATAA  | AAATTCAATC  | CTTTATGCTA  | TCGCTGTAGG  | CTCATTAACT | TGTTACATAT  | 360  |
|    | AATTCTTAAC  | TATCCTTTGA  | TGATTGTTTT  | ATTGAGTTGT  | TTGTTGATG  | GATACCTTCA  | 420  |
|    | CGAATTTCTA  | TAGTTCAATG  | CTACTAAAAA  | AACAGCCCAA  | AACCTTAAAT | TGTTTTGGAC  | 480  |
| 10 | TGTTTTATAA  | TTATGCTTGC  | GATGGTGTTT  | TAGTTTCTGA  | AGTTTGTTC  | GCAATGTCAT  | 540  |
|    | ATTTAACTC   | TTTACCATCA  | TGATCTACTG  | TAACCTTCTT  | ACCTTCAATT | TGATTACCAT  | 600  |
|    | CTAATATATA  | TTCACTTAAA  | TTATCTTCGA  | TAGTTTTTTG  | TATCGCTCTA | ATTAATGGTC  | 660  |
| 15 | TTGCACCATA  | TTCTGGATCA  | TATCCTTCTT  | CTGCGATTTT  | GTCTTTGCTT | TTATCAGTTA  | 720  |
|    | CAATAATATT  | TATGTTTTGT  | TCAGATAATC  | GATTTGTTAA  | TTTATTAACC | ATCATTGTTA  | 780  |
| 20 | CAATTCTCTT  | TAATTCTTCT  | TTTGTTAGTT  | TATGGAATAC  | AATGATATCA | TCTACACGGT  | 840  |
|    | TTAAAAATTC  | TGGACGGAAT  | GAATTTTTTA  | ATTCTTTTAA  | CATCGTTTTT | CGAATTGTTT  | 900  |
|    | CATAATCTTG  | TCCATCACTT  | GAACCACCGA  | ATCCAGCAAA  | TCGTTGATCT | TGTAATCTCT  | 960  |
| 25 | GTGCCCCAAC  | GTTTGATGTC  | ATTATGATAA  | TTGTATTCTT  | GAAATCAACT | GTACGTCCTT  | 1020 |
|    | TTGTATCTGT  | CAAATGTCCA  | TCATCTAAAA  | CTTGTAAATG  | AATATTAAAT | ACATCTGGAT  | 1080 |
|    | GAGCTTTTTT  | AATTTTCATCA | AATAAAAATTA | CAGAATATGG  | TTTACGTCTA | ACTTTTTTCAG | 1140 |
| 30 | TTAATTGTCC  | ACCATCATCA  | TGACCAACAT  | ATCCTGGAGG  | AGCACCAACT | AATCGGCTCA  | 1200 |
|    | CTGCGTGT    | TTCCATAAAT  | TCATCTATGT  | CTACACGGAT  | CATCGCATCA | TCATCGCCAA  | 1260 |
| 35 | ACATTGATTC  | AGCTAAAGCT  | CTAGCTAATT  | CAGTTTTTACC | AACACCAGTT | GGTCCAAGGA  | 1320 |
|    | AGATAAAGCT  | ACCAATTGGT  | CGTTTAGGAT  | CTTTTAACCC  | TGCACGGGCA | CGTCTAACCG  | 1380 |
|    | CTTTACTGAT  | TGAATTAAAC  | GCATCTTTTT  | GCCCCAATAAC | TCTCTCATGT | AATGTATCTT  | 1440 |
| 40 | CTAGACTAAG  | AAGTTTTTCA  | GATTCTGTTT  | CAITGATTTT  | AGTTAATGGG | ATACCTGTCC  | 1500 |
|    | ATCCTGCAAT  | AACTTCAGCA  | ATATCTTCTT  | CTGACAAATGA | AGTTGaCATG | CCAATTTGTG  | 1560 |
|    | CATTCTTCCA  | TTCATTTTTA  | GCTTCTTCAT  | ATTGCTTTTT  | AAGTTTTTGT | TGTTTTATCAC | 1620 |
| 45 | GCAGgTTAGC  | AGCATTTTCA  | AACTCTTGAG  | CATGACTCTG  | GGCATCTTTT | TCATTTTTTAA | 1680 |
|    | CTTTTTCAAT  | TTCTTGTTCA  | ATTTCTTTTA  | AATTATTAGG  | TGTCGTATGA | CTCTTAAAGTC | 1740 |
| 50 | TTACTTTAGA  | ACTTGCTTCA  | TCAATTAAT   | CAATTGCTTT  | ATCTGGTAA  | AAACGATCTG  | 1800 |
|    | AAACGTATCT  | GTTACTTAAAT | TTAACAGCTG  | CTTCAATAGC  | TTGTCGTGAA | ATATTAATAC  | 1860 |
|    | GATGGTGTC   | TTGTAACGA   | TCTCTTAATC  | CTTTTAAAAAT | AGCAACTGTA | TCTACTACTG  | 1920 |

|    |  |      |
|----|--|------|
|    | TTTTGGGATA TTCATCTAAT GTAGTAGCAC CAATACATTG TAATTCACCA CGTGCTAATG  | 2040 |
|    | CGGCTTCAA AATATTCGAA GCATCGATAG CACCTTCAGC ACCACCAGCA CCAACTAAAG   | 2100 |
| 5  | TATGCAACTC ATCAATAAAT AGGATGACAT TACCTGCTTG TTGGATTCTT TCCATAACCT  | 2160 |
|    | TTTTCAGACG CTCTTCAAAT TCACCAAGAT ATTTAGTACC TGCAACTACT GTTCCCATAT  | 2220 |
| 10 | CTAAAGACAT AACACGCTTA TCTTTTAATG TCTCTGGTAC CTCATTATTC ACTAAGGCTT  | 2280 |
|    | GGCTAAACC TTCAGCAATA GCAGTTTAC CAACACCTGG CTCTCCAATA AGCAGGAT      | 2340 |
|    | TGTTTTCTGT ACGTCTACTT AATACTTCAA TTACACGTGT AATTTCTTTA TCACGTCTTA  | 2400 |
| 15 | TAACAGGATC TAATGTACCG TCTTTGGCAA TGACTGTTAA GTCAAGAGCT AAACATATCTA | 2460 |
|    | AAGTGGAGT ATTATTTGAC TTACTAGCTT GTGCATTTTT ATTACTCATT TCAGGGTTTC   | 2520 |
|    | CTAAAGCTTT CACAACCTGT GCACGTGCTT TAGTAATATT TAAATCTAGA TTTGCAAAAA  | 2580 |
| 20 | CTCTTGCTGC AACACCTTCA TTTTCACGAA TCAAGCCTAA TAAATATGT TCCGTTCCAA   | 2640 |
|    | CAAAATTTGT ATGTAATTTT CTAGCTTCAT CCAATGATAA TTCAATGACT TTTTtagctc  | 2700 |
|    | TAGGTGTATA ATGCAATGTA CCAACATGAT CTGACCATG TCGATTAAAT TTTTCAACTT   | 2760 |
| 25 | CTTCAATTAC TTTATCTTCA GTGATATTAA AACTTTCTAA TACTTTTGCA GCAATTCCCTT | 2820 |
|    | CAGGTTCCTT CATTAACCCC AATAATAGGT GTTCTGTCC TATATTTCGA TGATTTAAAC   | 2880 |
| 30 | GAATTGCTTC TTCTGGGCA TGIGCTAATA CGCGCTGTGC ACGCTCAGTT AATCTACCAA   | 2940 |
|    | ATAACATAAA TAATGACCTC CTACTTTATA TGTTCTCTTA GTATATCTGC TCGTITTTCT  | 3000 |
|    | TTTACAGATT TGTCATCTTC TTCACTAAT AAAAATGGTG ACTGTATAGC TACCATCAAT   | 3060 |
| 35 | TCATTAAATT TAAAGTTTGG TAATTCGAAT TAATTTAAAT CTATACCAAG TTTAACTGCG  | 3120 |
|    | TTAATCTATA AGAAGCCTCT TCCATAGTTA TCATTCTACA GTTTTGTAAG ATACCTAGCG  | 3180 |
|    | AGCGAAAAAC ACGGCTTGT GTTTCTAATT GATTATAAGT GTCTAACTTT TGTCGTATTT   | 3240 |
| 40 | GTITTTCTTC ATGAATGATT TGATTAACAA CTTCGTGTTA TGTTCTATG ATTTCTAACT   | 3300 |
|    | CAGATTACC AAGTGAAGT TGGTTGGATA CTGATAAGT ATGTCATAA ACTTGCGAAC      | 3360 |
|    | CTTACCGTA AATACCTCTG ATTGTATATC CAAAACGATT AATGTTTGA GCAATCCGTG    | 3420 |
| 45 | TCATTCTTTT CATAATAGAT AGACCTGGCA AATGTAGCAT CACGCTGTCT CTCATACCAG  | 3480 |
|    | TACCTATATT GGTAGGACAT GTAGTTAAAT AACCAGTTG TTCATCATAA CTTATATCAA   | 3540 |
| 50 | GGCTTGATC TAATTCATCA TCAATTGATG AAGCTTGATT ATATAAGCC TGTAAATGTCG   | 3600 |
|    | TGTCAGTTCC CATAGCTTGA ATACGAATAT GGTCTCTCTT ATTTATCATG ACACCTAAAG  | 3660 |
|    | ATTCATGCTC ATTCACTAAT ACTGCGGCTG CTGGTTGTTT TATTAGTTCA GGACTAATCA  | 3720 |



|    |   |      |
|----|---|------|
|    | CAAGTTTGG CAAGGCATCT TGTACCTCAT TTATAACTCT AAATCCGATC ATTTTCAGTA  | 3840 |
|    | GCATACATTA GTGGATGCAC ATGATTTTCT AAATACGGC CTAACCGAAT TCTAGAAGAC  | 3900 |
| 5  | ATAACAATTG GTGTTTCTTC ATTACTTTTC ATCCATTGGC TGATATTATC ATGAATATTA | 3960 |
|    | TGGTCATCA TGTTCACCT CACTCTCAGC TTTTAGTGCT TTAATTTTCA CTCTACAAT    | 4020 |
|    | GGCTGCTTC TCAAAATCTT GGATTTCAT AAGTTTTTTC AAATATTCA TCTTTCTTC     | 4080 |
| 10 | GATTTTTCG TTTAAAGCTA TCTTTTATG TGAAGAATGT GGTGTCTTTC CAACGTGCTC   | 4140 |
|    | AAATGTCCA CCTTGAAGCTC TGGGACGAT ATCAATGATG TCATCTTTAA ATGTTGcATA  | 4200 |
| 15 | ACAATTAGCA CACCCAAATT TACCAACATG TGCAATATC                        | 4239 |

## (2) INFORMATION FOR SEQ ID NO: 334:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1245 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 334:

|    |  |     |
|----|--|-----|
|    | CAAACTTTG ATTATGCTG GTCAGTTATT AATGCATACA GTAATGCATC GTTGTATTAT    | 60  |
|    | TTTATTATT TACTATTAT TATGTACCAA ATCAGCCTTt CAGTAAAAATA GGTCAAATTA   | 120 |
| 30 | CTGATTTTCT AAATATAAAA TGCCCTCCTAA TAACATACTA TTAGTACATC ATTAAGAGGC | 180 |
|    | TCTTGTTTA TTTGCATCT AAGCGCTCAA ATTAAATTT AAGATGAAGA TTCTTGCAAT     | 240 |
| 35 | AATATTCTA TATTGTGCA TATATCCAAT GGATCTTCCA TTGGCAAGAA TCGATTTACA    | 300 |
|    | ACATTTCCTT GTCGATCGAT TATAAATTTT GTGAAATTCG ATTTTATTGG GgACCAAG    | 360 |
|    | ATTCTGGTT GTTCaTTCTT TAAATGCGTA AATAACGGAT GTTCATGTTT CCGTTTACA    | 420 |
| 40 | GATATTTTAG CTAACACTGG AATGTTTACA CCAAAATTCT CACGACTAAT TTtCAAGATT  | 480 |
|    | TCTTCATTAG AACCTGGTTG TCGATTGTCA AAATTATTAT TCGGAAAACT CAACACTACA  | 540 |
|    | AACCCACGAT CcTTATATTT TTGAAATAAA GTCTCTAGTT TTTTCAATTG TTGCTATAT   | 600 |
| 45 | ATACATTCTG TTGAGTATT AACAACTAAA ATCACTTTAC CTCTAAATGC TTCTAATTTA   | 660 |
|    | TAAGTTAAGC CTTTATAATC ACTTACTTCG ATATCATACA CATTTCtATT ATTcATAAGA  | 720 |
| 50 | CACCCCTACA CAGCCTTTTT TATATTGAAT ATGTTCTTTT TAGAATGTTc TGATAAAATA  | 780 |
|    | AGTGGCGGTT TACACGTGA ACACACATTA TATAGCGTGA TACATTTTTC GAGCACACGA   | 840 |
| 55 | TAAATAATGT TCGAGTTTAT GTTGTGCTC AACCTATCCG ATTTACCGTC TTTTTCACC    | 900 |

CCGTATAAC CCCTCATAT AATCATCCTT ATTTCTTATT TTAAAAAGA CAATTAGACC 1020  
 5 GCTCTTTAAA CTATAGAATA ATACTTAAGK TAAACTCATA CATACTGATA CCATACGTTA 1080  
 GATCTAACAA TTTAAAAATC GTTATAACTA TGGATTAAAG AGCTGCCCAA CTCATATAAT 1140  
 CCTTAAAAAC TTCACATGTG ATTGKtATT AAGCCCTCCT TTATCnTATT AAATATCCTT 1200  
 10 ATAACCCCTT TAAATTAATA CTGACACACT CATACTTGT TACAC 1245

(2) INFORMATION FOR SEQ ID NO: 335:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1399 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 335:

CTTGTAAGG TAGTTGTCA TTTAAATTA AACAGTATGC TAAGTThGTG CTTATATTTT 60  
 GCAACTTCAA TTCGACAGGC TCATCTTGAT ACATGAATGC CTCAAITTCa TCATGTGATA 120  
 25 ACTTTTTCAA AATATCAAIT AAATGrATAC TAAAAACGAT AAATAACATA TCCCAITTTG 180  
 AATTCAGCCC TAGCGATTTT AATTGTGTTA AAATCTTATC TTTTGA AAAA ATTCGATTCC 240  
 TAATGTCTTG TATATCGTTA TCAGTCAAAG TTCCCAATC TATAGTGAA TGAAGACCTA 300  
 30 AATAACACTT ATCCATTAA TATTCAATATA CCGTTAATGC AGAGACATCG AAACAACGTT 360  
 CTTCACTTAa AAAAACGCCA TCAACATCAa ATAAAAITTT CTTCACAATC CCCACTCCAT 420  
 35 TTCTGaAAAT TCAGaTATAA ATCATTCTAC TATTTGACTA AAAAAAGCGC AAACCCITATT 480  
 GAAGTaGATT TGCCTTTTAG CTGTThAAAT TTTATAAATG ThTTTCAAIT TCATCAGCAA 540  
 CCTGCTGTAC GTGTGTACCG ACAATAACTT GAGTTGAATG TTTGCCATTA ACAGTAACAC 600  
 40 CAACTGCACC GCGCTTTTTA ATCTTCTGTT TAICAATAAT AGATGTGTCT TTTAACTCTA 660  
 GACGCAACCT TGTTGCACAA TTGGTTAAAT TAACAATATT CTCTTGACCG CCTAAACCTT 720  
 CTAATATTTG TATAGCATGT TGATGATATT TACTTTGTTT AATAICATTT TCACAGGAG 780  
 45 CAATATTATC TTTTACAAC TGTGGGTCAA CTAATTCATT TTCACCICTA CCAATCGTAT 840  
 TCAAGTTAAA TACTTGGATT ACTACACGGA AAATACATA GTATAAGATG AAAAATACAA 900  
 50 CACCTTGAAC AAGCAACATC AATGGATGAT TTGACTATGG ATTAATTAGT GATAACACAT 960  
 AATCTATCAA ACCTGCACTA AATGAAAATC CAGCTGTCCA ATGGAATGTA GCTGCGATAA 1020  
 ATAAGATAA TCCTGTTAAT AACGCATGAA CAACATATAA GATTGGCACA ACAACACATA 1080

ACCAACCGTA AACTTGTTTT TTCTGAGTAG TTTTAGCTGT ATGATACATT GcTAACGCAG 1200  
 CCGCTGGAAT ACCGAACATC ATGATTGGGA AGAATCCCGC TTGATAGCGT CCTGTAAATAC 1250  
 5 CTTTATAGC ATCTTTGCCA CTTTGGAATT TACCAATATC ATTAATACCA ATCGTATCAA 1320  
 ACCAGAACAC ACTATTCACT GCATGATGTA ATCTGTAGG AATTAATAAT CTATTGGCAA 1380  
 CACCATATAT GAAAGCTCC 1399

## (2) INFORMATION FOR SEQ ID NO: 336:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1329 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 336:

TATAGTTTATA TTATTTAGCG AAGCATTAAAT ACTACCACCG GTTATAACAA ATGTATTTTG 60  
 CGTATTAAAT TGAATGGTAG GACCAATCAA AGTATATTCA ATCGCTGGAC CATCATTTGTT 120  
 25 AATTAAATGAC TGCGCACCT TAAAACTAAA TTGATCCATG GCACCTGCGC CTGAAAATCC 180  
 AATATGTTC TAACCTATTTC TTCTAGATC TTGTACCGIT GAAAAGAGAC CTGGTTGTAA 240  
 AATCTTAAAT GACATTTTCA ATCACCACCC AGTCATCAAC ATTAAAGTTG CCATCTGATA 300  
 30 TATCTCTTTC GATTGTGATA AATTTCTGTT CATCTATTGC ATAAAAATGT ATCCATTCTC 360  
 CTGCTTCGTA CATGACATT GGTTCACGCT CGCTGCTAAA TACTTTTAAc GGTGTGCGTC 420  
 CAATAATTTC CCATCCGCCA GGAGAATCTG ATGATATAG TCCTGTTTGA TTATTGCGAA 480  
 35 TACCTACAGA ACCTGCATGA ATTTTAAACC TTGGCTGATT ACGTCTAGGT GTATGTAGTT 540  
 GTTCATCAAG TcGCCTAAG TATGAAAATC CTGGCATAAA TCCTAGCATA TATATTTAAAT 600  
 40 AAGGTTTACT TGTATGTTTT TCAATAACTT GCTCAACAGT TATTGATTA TGCTTTGCTA 660  
 CTCTCTCAAT ATCTGGTCCA TATGTACCAC CATATTGAAC AGGTATTTTA ATAATACGAT 720  
 TGGTTTGATT CACAGCATGA ACATTTTTTT CATTAAATTT GTTAAGTTCT AAATTTTCAA 780  
 45 TTAATTTAGA AGATGTTATA GCTTGTTTAT CAAAATATAT TAGAACTGCT CGATACGAAG 840  
 GGACAATATC TTGAATTTCT AATATTTCTT TTCTCGTAT CCACCGTACC ATTGCTGTGA 900  
 50 CATACGATA TGCTCTTTCG GATATTTTAT TTCAAATAA AATCATAATT GTCTGCTCGT 960  
 TAATAAATCT TACATCCACT TTAATCCCC CTTTGATTG CAATAAACCA GTATTGAATA 1020  
 CCTTTTCATT GTATCATTTGA GAAGCACAAG TTGTTTAAATA AGTAATTCAA ATCGCATATA 1080

TTAATATTGT AACTCTTACA CTAATTTAGG TTCTGCTATC ATTCGGTCTG ATGGAAAATT 1200  
 TTTACTTTTC ATCTGTCCGA TTTTTTGATT TTGAATATAA AAAAGCACGA CCGAAGTATC 1260  
 5 ATTAACACAC TTCAATCGCG CAATTAAATA ATCTATTGTA TCATTTATTG GATATTAAAC 1320  
 ATTTTTCAG 1329

(2) INFORMATION FOR SEQ ID NO: 337:

10 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 3421 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 15 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 337:

20 GAATAGTCC ATTGGAAAG ACATAAGAA TCCAGAAGCA CCTTTAGCGA ATATGCCGTG 60  
 TTCATATAGA TTTGTGAAAC TGGCATGCCC GAATTGAGTT TTAATGCAA ATAGAATCAT 120  
 GACGAACCA ACTACTATTA AACCAATAAT TGTCGTATT TTAATGATAG AGAACCAAAA 180  
 25 TTCTAATTCT CCGAAAAGTC TTGCGCTAAG TAGGTTGAAT GACATTAAATA ACAGTACACA 240  
 AAATAGTGCA CTTATCCAGT TTGGAATTTT TGGGAACCAA AAGCTAACAT ATTTTGCCAC 300  
 AGCGGTACT TCAGCCATAC CTGTAATAAT CCAACAGAAC CAGTATGTCC ATCCGGTAAC 360  
 30 AAATCCTGCA AAAGGCCCAA TATATGTATT GGTACATCT GCGAAGATT TAAATTCAGT 420  
 ATTCTGTATA ATGATTCTCT CTAACCTCG CATAACATA AATAACATA ATCCTATAAT 480  
 GATGTATGTT AATAGAATTG AAGGGCCGGT TAATGCAATC GTTTGACCAG CACCTAAGAA 540  
 35 TAAGCCTGTA CCAATTGCAC CGCCAATTGC AATTAAATTG ATGTGGCGAT TGCTCAGTTC 600  
 CCTTGTAAT TTTTCAGCCA TAATACATCT CCTTAAATA TAGATATGTT TATTATGCAC 660  
 40 TTATATTGAG ATATATACAA TTATTTTCGG TAAAAATGTG TAAATTCCTA TGTTAATATA 720  
 CTTTGGTTTT TATAATCATA TATAATAACC AATTGAAAAA TTAATTCTAT TGTAATATTC 780  
 ATGATTATT CACATCTTGA AAAAGCTTTA ATGGTGCTAT TTGTGGCTAT TCTGTGACAT 840  
 45 TTACATAGAT TTACAAAAA ATTGTTGCAC ATATAATGCC AGTCTTTATA TTTCACAAAC 900  
 GAAATCGGTT TACTATAATA TTAGTTGAAA GCCATTTCAT AAAGAAACAG TAAAGGGGAA 960  
 50 ATTATCATA GCGaATTAC AAAGAGGTTT AGAAGGGGTT ATCGCGCGG AGACTAAAAT 1020  
 AAGTTCAATT ATTGAAGTC AATTGACTTA TGCCGGGCTAT GATATTGATG ATCTAGCTGA 1080  
 AAATGCGCAA TTTGAAGAAG TTATTTTCCT ATTATGGAAC TATAGATTGC CAAACGAAGA 1140

55

|    |   |      |
|----|---|------|
|    | TACACATTTT GAGGAGTATG TTACAGATCA CGTGCAATCCA ATGACAGCAT TACGTACGTC  | 1260 |
|    | ATTATCATAT ATTGCACATT TCGATCCTGA TGCTGAAAAA GAATCAGATG AAAATCGTTA   | 1320 |
| 5  | TGAAAGAGCA ATGCGTATAC AGGCTAAAGT AGCATCATTA GTTACAGCGT TTGCTCGAGT   | 1380 |
|    | AAGACAAGAT AAGAACCAC TTAAGCCTAA TCCTGACCTA AGTTATGCGG CAAACTTCCT    | 1440 |
| 10 | ATATATGTTA CGTGGGGAAT TACCAACAGA TATAGAAGTA GAAGCCTTCA ATAAAGCACT   | 1500 |
|    | TATTTTACAC GCTGATCATG AGTTGAACGC ATCTGCATT ACAGCAGGTT GTGCGGTATC    | 1560 |
|    | ATCATGTGCA GATATGTACT CAGGTATTGT AGCAGCCGTA GtTCTCTGAA AGGGCATTAT   | 1620 |
| 15 | CATGGTGGTG CAAACGAACA AGTTATGACG ATGTTATCTG aGATTGGGTC AaTTGAAAAT   | 1680 |
|    | GTTGATGCTT ACTTAGATGA AAAATTTGCT AATAAAGrTA AAGTAATGGG CTTCGGTCAT   | 1740 |
|    | CGTGTATATA AAGATGGTGm tCCTAGaGCG AaATATTTaA GaGAAaTGAG CGGTCAAaTT   | 1800 |
| 20 | mCGAAGAGC CTGGTCGTGA AGAATTATTT GAAaTGTCAG TGAAAaTGGA AAaMCGTATG    | 1860 |
|    | GCAGAGAAA AAGGATTAA TCCTAATGTT GATTTTTATA GTGCGAGTGT TTATCACTGT     | 1920 |
|    | ATGGAATAAC CTCATGACTT ATTCACGCCA ATCTTTGCTG TAAGTCGTTC TGCAGGATGG   | 1980 |
| 25 | ATTGCTCATA TTTTGAACA ATATAAGAT AATAGAATTA TGCGTCCTAG AGCGAAATAT     | 2040 |
|    | ATTGGCGAAA CGAATCGTAA GTATATCCCG CTTGfAGaAA GAAAmTAATC AATACAAATT   | 2100 |
| 30 | AAAAATGAAG ATGTAAAATT TGGAGGTAAA ATAACATGA CTGCAGAAAA AATTACTCAA    | 2160 |
|    | GGAACCTGAAG GATTAAACGT ACCTAATGAA CCAATTATCC CATTATTATAT CGGTGATGGA | 2220 |
|    | ATTGGACCGG ATATTTGGAA GGCAGCAAGC CGAGTTATAG ATGCTGCTGT TGAGAAAsCC   | 2280 |
| 35 | TATAATGGCG AAAAACGCaT TGAATSGAAA GAAGTGCTAG CTGGCCAAAA AGCATTTGAT   | 2340 |
|    | ACAACCTGSGT AATGGTTACC TCAAGAAAAA CTTGATACAA TTAAAGAATA TTTAATTGCT  | 2400 |
|    | GTTAAAGGAC CTTTAAACAAC ACCAATTGGT GGTGGTATTA GATCATTAAA TGTGGCTTTA  | 2460 |
| 40 | CGCAAGAAAT TAGATTTATT TACTTGCTTA AGACCGGTAC GTTGGTTTAA AGGAGTACCA   | 2520 |
|    | TCACCTGTTA AACGTCACA AGATGTTGAT ATGGTTATTT TCCGTGAAAA TACTGAAGAC    | 2580 |
|    | ATTTATGCTG GTATTGAATT TAAAGAAGGT ACAACAGAAG TTAaaaaAGGT AATTGACTTC  | 2640 |
| 45 | TTACAAAACG AAATGGGTGC GACAAACATT CGATTCCCAG AAaCTTCAGG TATTGGTATT   | 2700 |
|    | AAACCAGTTT CTAAGAAGG AACTGAGCGA TTAGTTAGAG CAGCTATACA ATATGCTATC    | 2760 |
| 50 | GATAATAACC GTAATCAGT TACTTTAGTT CATAAAGGTA ATATTATGAA ATTTACAGAA    | 2820 |
|    | GGCTCATTTA AGCAGTGGGG TTACGATTTA GCATTATCTG AATTTGGTGA TCAAGTATTTC  | 2880 |
|    | ACTTGCAAC AATATGACGA AATTGTTGAA AATGAAGGCA GAGATGCTGC TAATGCTGCT    | 2940 |

55

EP 0 786 519 A2

|    |  |      |
|----|--|------|
|    | TTACAACAAA TTTAACTCG TCCAGCTGAG CATGATGTTG TAGCAACTAT GAACTTGAAT   | 3060 |
|    | GGTGACTATA TTTAGATGC TTTAGCTGCA CAAGTTGGTG GTATTGGTAT TGCGCCAGGT   | 3120 |
| 5  | GCAACACATTA ATTATGAAAC AGGTCATGCT ATTTTGAAG CAACACATGG TACAGCTCCA  | 3180 |
|    | AAATATGCAG GTTTAAATAA AGTGAATCCA TCTTCAGTAA TTTTAAGTTC TGTATTAATG  | 3240 |
|    | TTAGAACATT TAGATGGCA AGAAGCGGCA GATAAGATTA CAGATTCAAT TGAAGATACA   | 3300 |
| 10 | ATTGCTTCAA AAGTTGGTAC TTATGACTTT GCCCGTTTAA TGGaTGGtGC TGAAGAAGTT  | 3360 |
|    | TCTACATCAG CATTTCAGAG TGAATTGATT GhaAAATTTAA AATAAGCAGA ATAGAATTAG | 3420 |
| 15 | G  | 3421 |

(2) INFORMATION FOR SEQ ID NO: 338:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3173 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 338:

|    |  |     |
|----|--|-----|
|    | CCCTnGATAC CCAAACTGCC CTAATACCAC TGGCTAAACC TAATGGATAG TACCATTTTAT | 60  |
|    | TTTCCAATAA ATAAACCAAC TGCAATnGCT ATaACTCTAA ATATAATAGA GATAATCmCA  | 120 |
| 30 | TTAATTGGAT TAATACGCCA AATACTAGTA ATAATAGGCT AGATAATAAT CCACCTAAAA  | 180 |
|    | AGTACTTTTT AATTCCAAGG AAAGCTAATA TCAATAATGC TGCCGGTGCA GATAATTGAA  | 240 |
| 35 | AATCTAATCC TGGTATAATG GACGGTATTT TCAAAACTGC CAAAATGGTT AAAATGCGAG  | 300 |
|    | CAATGACACT AATTTGAGTA ATATCTTTTG ATGTCACTAT AAAACCCCTA TACCGTTTCA  | 360 |
|    | TAAACAACCTT GCTTCGGTGT GCTTTCTAAA AATGATATGT AATGATTTAA ATCAATACAA | 420 |
| 40 | TCGTCCACAA ATATTATTCT GCCTCCATAT CTCGTATTAA CTGGTTTAAT ATCAAATAAT  | 480 |
|    | CGATGGTAAC CAATTTTAGC AGCGGCAAAA TAACCTGTGC TATACGTTAA GTCATCGGAC  | 540 |
|    | ACGCAAAAGT CTCCTTTGAC ATACGGATGC GCATTGATC AACTAGCAAT TGCTAAGGCA   | 600 |
| 45 | TCAGTCACCT TTTCAATTAAG ATCACCTTTA TTATTTATAT CTTCAAACGA AAAATGTGTT | 660 |
|    | GCCCTAATCC CCCTTTGTC AAATGAATCT AAACGTTTAC CAGATATAGC AGATAGAATA   | 720 |
| 50 | ATAGTCCCTG TATAAACCGT TTCAITTTTA ATATATGTCA TCCCTTGATT TAGCGCTGT   | 780 |
|    | TCAGTGACAC CACATTCTTG TGTAAATGT TGGAGATTG CTTTATCATC CTCAATAATT    | 840 |
|    | TGTAATGCTT TTATTTGTG AATCGGTTCC ATGATTTTTT GTATTTTTAT ATTGAGAAAA   | 900 |

|    |            |            |            |            |            |            |      |
|----|------------|------------|------------|------------|------------|------------|------|
|    | ATTTTGTCAA | ACTCACAAT  | CGTTTCAGCA | CCACTAATAT | GAACATCTTG | ATTGCTAGAA | 1020 |
|    | CGCATTTTTA | TACTATACAT | GACGATCACC | TcAATCTTCT | TGaTGCAAAA | TTTCAACAA  | 1080 |
| 5  | CCTATCTATA | TCTTGTTT   | CAG        | TATGAAAATA | CGACAATGAT | ATTCTTAACA | 1140 |
|    | CACAGTGGGA | TACCTTAAAT | AACTTGTA   | AA         | CACATGATGC | TTTAATAATG | 1200 |
|    | GTTCTCAGCC | GCTTCTATGT | CATCAAACTC | AATAAACTTA | ATCGCGAGST | TTGAACTATT | 1260 |
| 10 | ATAATKAACA | TTGAGTGT   | TTAACTTTTG | GTTAAATAT  | TTACTCAAAC | TATTTAATTT | 1320 |
|    | AGTGCCTCTA | TCATCAGCAT | TTATTAACCT | TTCAATGTTT | CTTTTATAAA | AATACAAAT  | 1380 |
| 15 | ATAAATGGC  | AAACTACTTG | AGTAGATGAG | TGGTCTACCG | TGATTAATTA | aCATATCCT  | 1440 |
|    | CaCATCaTTT | GaACTKaAAA | TcACACCCCC | GTATGCACCA | CATGCTTTAG | ATAAACTAGA | 1500 |
|    | AGTGAGTATA | TCTACACCTT | GATAATTGCA | GTAATCTCT  | ATcCCAAAC  | TATGTAAAC  | 1560 |
| 20 | ATCGAGTATC | AGTGTGCGT  | TAAATTTATG | CTTTAATGAG | ACTAATTGAC | CAATATCCAC | 1620 |
|    | AACGTGCCCA | TTCGTTGAAA | ATACACTATC | AGATATGATT | ATTTTGGTA  | TATTTTGATT | 1680 |
|    | AGGGTATTTT | TCTAACCTTT | TTTCTAAATC | AGCAATATCT | AAATGCTTAT | ATATCACTTT | 1740 |
| 25 | TTCTAAACCA | CTTAACTTAA | TACCGTCAAT | AATACTCGCA | TGATTTTCTT | GATCTGAAAA | 1800 |
|    | CAGCACACAA | TTTGATTTTT | TGAAAAATAT | AAATAACGCC | AAATTAGCAT | CATAACCACT | 1860 |
| 30 | ATTTAAGATA | GTACATGcAC | TATATCCGAG | CCAACCTGCT | AACATTGTTT | CAATTTCTTC | 1920 |
|    | ATAAGCTGTC | GAACTCCAC  | TAATTAATCT | TGAACCTGAT | AAGTGATAAC | TATACTTCCG | 1980 |
|    | CATAAATCTT | TCGAAATCAT | CCTTATCAAA | CGCTAATTGA | CCTAATCCTA | AATAATCAT  | 2040 |
|    | AGATGTATAG | TTCGTACATC | TCTTATTTTC | TACTTCAATA | TACTGTCTAT | CTATATACCC | 2100 |
| 35 | TACCGATTTA | AGCGACCGAT | ATAACCTTTT | CTGTTGTAAT | AAATCAATTT | GCTCTGAAA  | 2160 |
|    | CTTCATICTT | GTTTTCCTTA | TTTTCACAAG | TGTCATAATC | AATTTCAAAG | CCTAAATCAT | 2220 |
| 40 | TAATCATATC | GTAGTCTAAT | TGGTTCGGTT | GCCCAACAGT | AATTAGATAA | TnCACCGACA | 2280 |
|    | AATATTGAAT | TCGCCGCTTT | TAATGCTAAT | GGCTGTAACG | AACGTAAGTT | GACCTCTCTT | 2340 |
|    | CCTCCAGCAA | TACGAATTTT | TTTCGTAGGA | TTGATTAATC | GGAATAATGC | TACGATTCTT | 2400 |
| 45 | AAACATTCTA | TTGGTGTTAA | ATCATCCATG | CTTCCAACT  | TTGTGCCTTT | GATTGGATGC | 2460 |
|    | AAAAAATTAA | TCGGAATACT | GTCCGCATCC | ATTCTTTTA  | AAGCAAATGC | CATATCAACA | 2520 |
|    | ATATCTTGAT | TAGATTCTCC | CATACCACAA | ATCACGCCAG | AACATGGTGA | TATATTATTC | 2580 |
| 50 | GCTTTCATTA | GTTCTATCGT | ATCTGTTCTA | TCTTTATAAC | TATGCGTTGT | CACGACGTTA | 2640 |
|    | TCATGGTAA  | TTTCACTTGT | ATTAATATTG | TGGTTATATC | TGCTACACCC | AGCTGACTTA | 2700 |

EP 0 786 519 A2

|    |  |      |
|----|--|------|
|    | TGTTGAGATT TAATCGTTCT TACAGTATTA CTAATATGAT CAACCTCTTT ATCGCTCGGT  | 2820 |
|    | CCTCTACCACT TCATAACAAT ACAATATGTT CCAATATGAT TATCATGTGC CACCTTTGCT | 2880 |
| 5  | CCATCGATAA TTGTCTCTC TGGAAATAAA GCATATCGTT GTTTTGTGTT AATATCTCGT   | 2940 |
|    | GATTGTCCAC AGTACCACCA ATTTTCAGGA CATATACCAC TTTTAGCATT TAAATCATG   | 3000 |
| 10 | TTTAATTTTA CTTTTTACC AAAATAATGT TTTCTAAAA TGTACGCCTC ATTTAATAAA    | 3060 |
|    | TCTAAGGTAT CAATATTAGT ATCTCATAA ATCTTCAATA CAGTCTCTTT TGTCAATTGT   | 3120 |
|    | CCCCCTTGTA ATATCGGTTT AGCCAAATTC ATATTAACAC TTCCTATCTA AAA         | 3173 |
| 15 | (2) INFORMATION FOR SEQ ID NO: 339:                                |      |
|    | (i) SEQUENCE CHARACTERISTICS:                                      |      |
|    | (A) LENGTH: 1694 base pairs  |      |
|    | (B) TYPE: nucleic acid   |      |
|    | (C) STRANDEDNESS: double   |      |
| 20 | (D) TOPOLOGY: linear   |      |
|    | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 339:                         |      |
| 25 | CGATTATCCA TTAATACAAC CCTAAGTAAA TGTATAAAAA TTATCTTCCA CAAACTCAA   | 60   |
|    | CAAAAGCCTA AATAAATTAC AGCAATTTAT CAAATATTGC TTACTTTGAT TTTATGAAAT  | 120  |
| 30 | nACTTAATTC TAACACATAC TAAATCATCA TATACTAATT CGAAATCAAA TGCATTAGA   | 180  |
|    | GATAATCaAA ATGCGGAAAC ATCTCCAATA ATCAATAATC TATTCCTAAT AAATATGAAT  | 240  |
|    | GTCTCAACA ATACATTATT TATATCTCTT TACACTGTCA TCGACAAAA CTAATCTTT     | 300  |
| 35 | CACCTTCAAT TTGGAACGTG GTTCTACGAC ATTTGCTGCT ATATCATTTA ATGGGATTAA  | 360  |
|    | AACAAATGCA CGTTCATTCA TTCTCGGATG TGGCACCGAC AGTTTTGGTA AATCTATCAT  | 420  |
|    | TTCTTCCA TACAACAAAA TATCCACATC TAAAGTTCTA GGACCCCATC GTTCCTTCT     | 480  |
| 40 | AATACGGTGT AAACATTCTT CTGTCTKCAA ACAACATTCC AACAGTTGTA ATACTGTGAG  | 540  |
|    | TGTTGTTTsA ATTTCAACAC ACAAAATTAA AAAGTTAGGT TGCTCAGTAT ACCCAACTGG  | 600  |
|    | TGCTGTTTCA TAAATCGGAG AAATAATAGA TACGTTAATA CCATCATATT CATTCAAAAT  | 660  |
| 45 | CTTyATAGCA TCGTTTAACT GGCTTTCTCT ATCACCaaTa TTACTACCTA AyCCTAAGTA  | 720  |
|    | TGCTTGAATC ATyTATTCTC CCTCACTATT TCGATACCTA CTCATCATA ATGACCCGGA   | 780  |
| 50 | ATCGGTGGGT TTTCTTTAGT GATTCTCACT TTCGTTTCCA TTACAGGATT ATATTGTGAA  | 840  |
|    | TTTATACGAT TTGCAATACG TTCAGCTAGA TGCTCAAGTA AATTAACGGC CTTACCTTCC  | 900  |
| 55 | ATAATTGATT TAACCTCTTc GAACACTTCA CCATAATGAA CTGTATCAAT AACATTATCA  | 960  |



EP 0 786 519 A2

ATTCATTTT CAGCTGATAA AGCACCATGA TATCCATAAA AGCGCATACC TTAAAGAAAG 1080  
 ATGTGTCTTT GCATTTTCAT TCTCCTTTAA AAAATCTATA CCTTTAGCTA ATTTAGCATT 1140  
 5 CAACTCGACA TTATGAACGC GTACTGCTCT AACGCCTTC ATAATACCAT ATGCAGTCGT 1200  
 AGCTGCAGTT ACTTCATCTC TTCAACCGG TGTGTATCA TAACCCATCA TCTCTTTAGT 1260  
 GAAACGTTTC CGGCTTGTCT CTAATAAAAC TGGATATTCT GTTGCAACAA GTTCATCCAG 1320  
 10 TCTTGCCATA ACTTCGGCTT CTCATTCTCT AGTTTTAGCG AAACCTATAC CTGAGCTAG 1380  
 CCAAAATTTTA TTGAAGGTA TACCAGCTAT TTTAGCTTGA TGTGCTTGTG CTAACAAGA 1440  
 TGTTAACATT TCITCGACAA CCGGTTTCATC ACGATTACCA TTTCCATTAT GCATTAATAA 1500  
 15 AATTCCGCGC TCATATTTAG CTACAATTG GAACATACGA TGATCATACA GACCGCCAT 1560  
 TGATCATTA TCATATCAAC GGCTAATTTT AAACATGCTT CAGCAACCTC ACTTCGAAAT 1620  
 20 GTATGACTG AATTTTTTTA CATCAAAACC GACATAGCT TCACCAACAG TAATACTCTG 1680  
 TTCCATCTCT TCTG 1694

(2) INFORMATION FOR SEQ ID NO: 340:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1358 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 340:

AGCATTTCTT TCTATAAAC ATTTAATTGA ACATTATTAA GTACACTATT ACTATAGTCA 60  
 35 CTATATTGAA CaCATACCTC ATTTAATTCT AATAGCGGTT CAGATTGTGA CTTATTATCA 120  
 TTATTTGCAG ATGTTTCATC TATCCATTTT TTCACITTTAA ATTTAACATG TTCACCTATA 180  
 40 CAAACGTCAC GTAAATTGCG TAAGTTATCA ATGGATTGGA CATCTACTTC TGCATATTTA 240  
 AGCGCTGTAC AGTATAATGG TTCACGTATG CCTGCTTCTT TAAGCTTAGA TGATTTTAGC 300  
 AAATCACTAG GCGTTGTATT AGCGATGATT TTTCCATCTT TAAAAAGAAG AACTCTATCA 360  
 45 AACGTATCAT CTAATGATTC TTCTAATCGA TGTTCGACAA TAATCATCGT TGACTTTGTT 420  
 TCTTCATGAA TATTGTTTAA CAATCTCAGC GTTTCATGTC CTGTCGCAGG ATCTAAATTG 480  
 GCCAGTGGCT CATCCAATAT TAAATAGGC GTACGATGGA TTAATATACC ACCTAATGAA 540  
 50 ACGCGTTGTT TTTGACCTCC AGATAAATCT TGCGGTCGGT GATTTAAATG TTCTATCATG 600  
 CCAACTTTTT CAGCCCAATA ACTTACATTT TTCTTCATAT CATCTTGTTT AACACAATTA 660

TCTTGTA AAA CTGTACCAAC AACATTAGAT CTATCATGTA AACAACTAAC GGTGTCATCT 780  
 TGATTATTIA TATATAGTTC CCCAGTTATG TTACCTTTAG TTTTAAATGG AATTAAATCCG 840  
 5 TTTATGCAAT TTGCAAAAGT CGATTTACCA CTACCCGAAG CACCAACTAC TAATACTTTT 900  
 TCTCCTGGAT AAATATCAAC ATTTATATTC TGTAAATGTAG GTGTTGCTTG ACTATGATAT 960  
 10 TGAAAACTAA AGTCTTTGAA CGAGATAATT GGTTCAGTCA TGATATATCA TTACCTTTCT 1020  
 ATATTCATTT ACATATCTGA TTCAACAAAA TAACTATTCC TTACGTAAAC TACCTTTTTT 1080  
 AATTTGAGAT GAAGCATATG CTTTTAATAA TATTGTCCCA ATAATGCCAA CTGAAATAAT 1140  
 15 ATTTAATACT GCAGAGATAA CACCTTGTGT ATAAACCTTG TTAGCTGGTT CGTTATAAAT 1200  
 CAAAAATCTC AATGTTGGTG CAATAAGTGC CCAGCAAATA ATATTCGCCA TAATTTGACC 1260  
 GATATTA AAA TAAACCATCG ATTTCTTAGA AAATCGGCCT GAAGAAAGAT TTAATTTTAG 1320  
 20 TCCAATCCAG CCATATAAAC AGCCTATAAT TCCCGAGC 1358

## (2) INFORMATION FOR SEQ ID NO: 341:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4557 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 341:

TAGAAGAAAT GGAGAAAATG CTAATTCAAT TGTCAACATT CCATAGTTAT CATGATTTAG 60  
 AGTTTCTATT TGTGACACGT GAAGATGAAG TTGAAACATT GAAATGGGCA CGTTGGTTGC 120  
 35 CACATATGAC ATTGAGAGGG CAAAACATTA GAGGATTTGT TTACAATCAA CGAAGCGGTG 180  
 ACCAAATTTT AACGTCAATT TATAGCATGA TTAAGAAGCG TATCCAAGCT GTGCTGAACG 240  
 40 CAGCAGAAGT AATGAGCAAA TTATTTTTCAC ACCGCAATTA GTGTTTGTCA TTACAGATAT 300  
 GTCATTAATT ATTGATCATG TCATTTTAGA ATATGTAAAC CAAGATTAT CAGAATATGG 360  
 TATTTCAATTA ATCTTTGTGT AAGATGtGAT TGaAGTTTG CCAGAGCATG TAGATACCAT 420  
 45 TATTGATATC AAGTCTCGTA CTGAAGGCGA ACTGATTAGC AAAGAAAAAG AATTAGTTCA 480  
 ATTGAAATTT ACACCTGAAA ATATTGTAA CGTCGATAAA GAATATATCG CGCGACGTTT 540  
 GGCGAATTTG ATACACGTCG AACATTTGAA AAATGCAATT CCTGATAGTA TTACATTTTT 600  
 50 AGAGATGTAT AACGTGAAG AAGTAGATCA GCTTGATGTG GTTAATCGAT GGAGACAAAA 660  
 CGAAACATAC AAAACGATGG CAGTACCTTT AGGTGTAAGA GGTAAAGATG ATATTTTATC 720

|    |  |      |
|----|--|------|
|    | AGGGAAATCT GAGATTATCC AATCATACAT TTTATCTTTA GCTATTAATI TTCACCCCTCA | 840  |
|    | TGAAGTTGCA TTCTATTGA TTGACTATAA AGGTGGGGGT ATGGCGAACT TATTTAAAGA   | 900  |
| 5  | TTTAGTCCAT TTAGTTGGTA CGATTACAAA CTTAGATGGC GATGAAGCGA TCGTGCCTT   | 960  |
|    | AACATCAATC AAAGCCGAAT TGAGAAAACG TCAACGTTTA TTCGGAGAGC ATGATGTTAA  | 1020 |
|    | CCATATTAAT CAATACCATA AGTTATTTAA AGAAGSTATT GCGACAGAAC CAATGCCACA  | 1080 |
| 10 | TTTATTCAAT ATTTCGGATG AGTTTGCCGA ATTAAAAACA GAACAACCTG ATTTTATGAA  | 1140 |
|    | AGAAGTTGTA TCAACGGCAC GTATTGGACG TTCGTTAGGT ATTCAATTAA TACTTGCAC   | 1200 |
| 15 | ACAAAAACCA TCGGGTGTG TTGATGACCA AATTTGGTCT AACTCTAAAT TTAAGTTGGC   | 1260 |
|    | ATTTAAAGTA CAAGATAGAC AAGACAGTAA TGAAATTTTA AAAACACCAG ATGCAGCAGA  | 1320 |
|    | CATTACMTTA CCGgTCGTG CGTATTTACA AGTTGGTAAT AATGAATTTT ATGAATTTAT   | 1380 |
| 20 | CCAATCTGCA TGGAGTGGTG CAACATATGA CATCGAAGGC GATAAATTAG AAGTTGAAGA  | 1440 |
|    | TAAAGAGATT TACATGATTA ATGACTATGG TCAACTTCAA GCAATCAACA AAGACTTGAG  | 1500 |
|    | TGGACTTGAA GATGAAGAAA CGAAAGAAAA TCAAACTGAG TTGAAGCGG TCAATAGATCA  | 1560 |
| 25 | TATCGAATCT ATTACAACAC GATTAGAAAT CGAAGAAGTT AAGCGTCCAT GGCTACCACC  | 1620 |
|    | ATTGCCAGAA AATGTATATC ATGAAGATTT AGTAGAAACA GATTTCAGAA AATTATGGTC  | 1680 |
|    | AGATGATGCA AAAGAAGTGG AATTAACTTT AGGACTTAAA GAGCTACCAG AAGAACAATA  | 1740 |
| 30 | TCAAGGACCG ATGGTATTGC AATTGAAAAA AGCTGGGCAC ATCGCGTTAA TCGGAAGTCC  | 1800 |
|    | AGGATATGGT AGAACAAAGT TCTTACACAA CATTATTTTC GATGTTGCAA GACACCATCG  | 1860 |
|    | TCCTGATCAA GCACACATGT ACTTGTTGTA TTTCGGTACC AATGGTTTGA TGCCAGTTAC  | 1920 |
| 35 | AGACATACCA CATGTGCGTG ATTACTTTAC AGTAGATCAA GAAGACAAGA TTGCTAAGGC  | 1980 |
|    | GATACGTATA TTTAATGATG AAATTGATCG TCGTAAGAAG ATTTTAAGTC AGTATCGTGT  | 2040 |
| 40 | CACTAGTATT TCTGAATATC GAAAATTAAC TGGTGAAACA ATTCCGCATG TCTTTATTCT  | 2100 |
|    | TATTGATAAC TTTGACGCAG TAAAAGATTG ACCTTTCCAA GAAGTTTTTG AAAATATGAT  | 2160 |
|    | GATTAAAGTG ACGCGTGAAG GGCTAGCATT AGACATGCAA GTAACCTTAA CTGCTTCAAG  | 2220 |
| 45 | AGCTAACGCT ATGAAAACAC CAATGTACAT TAATATGAAA ACGCGTATCG CAATGTTTTT  | 2280 |
|    | ATATGATAAA TCAGAGGTGT CGAACGTAGT AGGACAGCAA AAATTTGCGG TTAAGATGT   | 2340 |
|    | TGTGGTCTGA GCATTGTTAA GTAGTGATGA CAACGTATCA TTCCATATTG GCCAACCAT   | 2400 |
| 50 | TAAACATGAT GAGACCAAT CATATAATGA TCAAAATTAAT GATGAAGTAT CGGCGATGAC  | 2460 |
|    | AGAATTTTAT AAAGGTGAAC ACCAAATGAT ATTCCTATGA TGCCAGATGA AATTAAATAT  | 2520 |

|    |             |             |             |             |              |             |      |
|----|-------------|-------------|-------------|-------------|--------------|-------------|------|
|    | GGAATTAGATT | ATGAAAGGTGT | TACACTACAA  | AAAAITTAAT  | TAACTGAACC   | AGCAATTGATT | 2640 |
|    | TCATCAGAAA  | ATCCGAGAGA  | AATTGCGCAT  | ATTGCTGAAA  | TTATGATGAA   | AGAAATTGAC  | 2700 |
| 5  | ATATTAAATG  | AAAAATATGC  | GATTTGTATC  | GCAGACTCAA  | GTGGAGAGTT   | TAAAGCTTAT  | 2760 |
|    | AGGCATCAAG  | TGGCTAACTT  | TGCCGAAGAA  | AGAGAAGACA  | TTAAAGCGAT   | TCATCAACTA  | 2820 |
|    | ATGATTGAAG  | ACTTAAAGCA  | AAGAGAAATG  | GACGGCCCAT  | TTGAAAAAGA   | TGCATTTTAT  | 2880 |
| 10 | ATTATCAATG  | ATTTTAAAAA  | ATTTATTGAT  | TGCACGTATA  | TTCCGGAAGA   | TGATGTTAAA  | 2940 |
|    | AAGCTTATTA  | CAAAAGGACC  | AAGAACTTGGC | TGACCATTTT  | TATTTGTCCG   | CATCTATAAA  | 3000 |
| 15 | GAATTAATAG  | ATGCTTATGA  | TAAACAGATT  | GATGTTCGAC  | GTA AAAATGAT | TAA CCAATTT | 3060 |
|    | AGTATAGSTA  | TTCTGATTTC  | AGACCAACAA  | TTCTTTAAAT  | TTAGATTTAT   | TCAACGAGAA  | 3120 |
|    | CCTGTTATTA  | AAGAAAAATGA | AGCATATATG  | GTGCGAAACC  | AAGCTTATCA   | AAAGATTAGA  | 3180 |
| 20 | TG GTTTAAAT | AGCAATGAAT  | TA AATAGGAG | GGAGGTATGT  | TATGAATTTT   | AATGATATTG  | 3240 |
|    | AAACAATGGT  | T AAGTCGAAA | TTTAAAGATA  | TTAAAAAGCA  | TGCTGAAGAG   | ATTGCGCATG  | 3300 |
|    | AAATTGAAGT  | TCGTTCTGGA  | TATTTAAGAA  | AAGCTGAACA  | ATATAAGCGA   | TTGAATTTTA  | 3360 |
| 25 | ATTTGAGTTT  | TGCAGTAGAT  | GATATTGAAA  | GCACAGCAAA  | GGACGTACAA   | ACTGCAAAAT  | 3420 |
|    | CTAGTGCTAA  | TAAGGACAGT  | GTA ACTGTTA | AGGGA AAGGC | GCCCAATACG   | TTATATATTG  | 3480 |
|    | AAAAAAGAAA  | TTTGATGAAA  | CAAAAGGTTT  | AAATGTGGGG  | TGAAGATATC   | GATAAAAAAT  | 3540 |
| 30 | AAGAACTCCT  | CCAAAAGACT  | AAGGAAGTAT  | CTGGCGGAAA  | GGCAAGGTAA   | TATTTTAAATA | 3600 |
|    | AAGCAATGAA  | TTTAATTTGA  | GGTGAAGATA  | TGGGTGAGTA  | TAAAGGTAAT   | AATGACAGATG | 3660 |
|    | GTGGCAAGGT  | TGATCAAGCG  | AAACAATTAG  | CGGCAAAAAC  | AGCTAAAGAT   | AITGAAGCAT  | 3720 |
| 35 | GTCAAAGCA   | AACGCAACAG  | CTCGCTGAGT  | ATATCGAAGG  | TAGTGATTGG   | G AAGACAGT  | 3780 |
|    | TCGCCAATAA  | GGTGAAGAGT  | GTGTTACTCA  | TTATGGCAAA  | GT TTCAAGAA  | GAATTAGTAC  | 3840 |
|    | AACCGATGGC  | TGACCATCAA  | AAAGCAATTG  | ATAACTTAAG  | TCAA AATCTA  | GCGAAATACG  | 3900 |
| 40 | ATACATTATC  | AATTAAGCAA  | GGGGTTGATA  | GGGTGAACCC  | ATGATGAAG    | ATGTTAAGCG  | 3960 |
|    | AATAGATTAT  | TTTTCTTACG  | AAGAATTAAC  | AATTTTAGTG  | GGTAGTAAAT   | TG CCTCTCGT | 4020 |
| 45 | AAATTTTGAA  | TTGTTTGATC  | CATCAAAATTT | TGAAGAAGCT  | AAAGCTGCTT   | TAATTGAAAA  | 4080 |
|    | GGAATTAGTA  | ACAGAGAATG  | ACAAGTTAAC  | TGATGCAGGT  | TTTAAAGTGG   | CTACATTTAGT | 4140 |
|    | CAGAGAGTAT  | ATTAGCGCCA  | TTGTAAATAT  | TGCA AATGAT | GATATGATGT   | TTGCACCATT  | 4200 |
| 50 | TAGAGTAGAA  | AAAGATGAAT  | ATATTTTGGT  | AAGCCGGTTT  | AAAAATGAATG  | GGTTTCAAAAT | 4260 |
|    | ACGAATTATC  | ATAAAGACGA  | TGCGATGGTG  | GTGCAATGTA  | CAACATATCT   | CTTTATTGAT  | 4320 |

CTTAATAAT GAAAGTATCG ATACGATTGG GCGTGTTTTA GAAATIGAAA TATACAATCA 4440  
 TCAAGGTGAC CCTCAACAAA GTTTATATAA CATTATGAA CAAAATGATT TGTTATTCAT 4500  
 TCGATACCCA TTAAGAGATA AAGTGCTGAA TGTTTCATATT GGTGTCATTA ATACATT 4557

(2) INFORMATION FOR SEQ ID NO: 342:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 3931 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 342:

TTGAGTGACT TTATTGAAGC GCGGTGTAGAA GAAATATTCT TCGAAGTATT TGATGTTTTA 60  
 CAAGATTTAG GATTAAACAA AGTAAATGGT GGGTTTATTG TAACGTGGTG ATCTGCAAAC 120  
 TTACTTGGCG TAAAGAATT ATTATCAGAT ATGGTAAGTG AAAAGTTAG AATTACACG 180  
 CCATCACAAA TGGGAATTAG AAAACCTGAA TTTTCTCAG CAATTCTAC AATTCTAGT 240  
 AGTATCGCTT TTGATGAGTT ATTAGATTAT GTTACAATTA ATTATCATGa TAATGAAGAA 300  
 ACTGAAGAAG ATGTTATTGA TGTGAAAGAC AAAGATAACG AATCTAAATT AGGCGgaTTT 360  
 GaTTGGTTTT AACGTAAAC AAACAAAAAA GATACTCATG aAAATGAAGT AGAGTCAACA 420  
 GATGAAGAAA TTTATCAATC AGAAGATAAT CATCAGGAAC ATAAACAGAA TCATGaACAT 480  
 GTTCAAGACA AAGATAAAGA TAAAGAAGAA AGTAAATTCA AAAAACTAAT GAAATCTCTA 540  
 TTTGAATGAT TAITGGCCAA TAAAACTAGG AGGAAATTTA AATGTTAGAA TTTGAACAAG 600  
 GATTTAATCA TTTAGCGACT TTAAAGGTCA TTGGGTAGG TGGTGGCGGT AACACGCGG 660  
 TAAACCGAAT GATTACCAC GGAATGAATA ATGTTGAATT TATCGCTATC AACACAGACG 720  
 GTCAAGCTTT AAACCTATCT AAAGCTGAAT CTAATAATCCA AATCGTGAA AAATTACAC 780  
 GTGGTTTAGG AGCAGGAGCT AATCCTGAAA TCGGTAAAA AGCTGCAGAG GAATCTCGTG 840  
 AACAAATTGA AGATGCAATC CAAGGTGCAG ACATGGTATT TGTTACTTCT GGTATGGGTG 900  
 GCGGAATCGG TACTGGTGCA GCACCAGTCG TTGCTAAAT TGCAAAAGAA ATGGGCGCAT 960  
 TAACGTITGG TGTTGTAAC TCGTCCATTTA GTTTGAAGG ACGTAAACGT CAAACTCAAG 1020  
 CTGCTGTGG AGTAGAAGCT ATGAAAAGCTG CAGTAGATAC ATTAATCGTT ATACCAATG 1080  
 ACCGTTTATT AGATATCGTT GACAAATCTA CGCCAATGAT GGAAGCATTT AAAGAAGCTG 1140  
 ACAACGTGTT ACGCCAAGGT GTACAAGGTA TCTCAGACTT AATCGCTGTT TCTGGTGAAG 1200

|    |  |      |
|----|--|------|
|    | GTATGTGGT TCTTCTGGT GAAAATAGAG CGGTAGAGC TGCTAAAAA GCAATCTCTT      | 1320 |
|    | CTCCATTACT TGAACATCT ATCGTGGTG CACAAGGTG GCTTATGAAT ATTACTGGT      | 1380 |
| 5  | GCGAGTCATT GTCATTATTT GAAGCACAAG AGGCTGCTGA TATTGTCCAA GATGCTGCAG  | 1440 |
|    | ATGAAGACGT TAATATGATT TTCGGTACAG TTATTAAATC TGAATTACAA GATGAGATTG  | 1500 |
|    | TTGTAAACGT TATTGCAACT GGTTTTGATG ACAAACCAAC ATCACATGGT CGTAAATCTG  | 1560 |
| 10 | GTAGCACTGG ATTCGGAACA AGCGTAAATA CTTCTAGCAA TGCAACTTCT AAAGATGAAT  | 1620 |
|    | CATTCACTTC AAATTCATCA AATGCACAAG CAACTGATAG TGTAAGTGAA AGAACACATA  | 1680 |
| 15 | CAACTAAAGA AGATGATATT CCTAGCTTCA TTAGAAATAG AGAAGAAAGA CGTTCAGAA   | 1740 |
|    | GAACAAGACG TTAATCGGTT AATATATATA CACAAATAAT TCAACACAAA TCATCAGATA  | 1800 |
|    | ACATATCTGA TGAATTTTTT ACTAATTTTT AGAACATGTA GAAGGACATT TAAGTTTTTC  | 1860 |
| 20 | AAAGTTATTA AAGTGTGTTA AGTATCGTGT GAAAATTAA TCAAAAAATTA TTTGCGCAAC  | 1920 |
|    | ATTTTAACTT TAAACATAAA TGTATATTA TATAATTATT AACTTTGTAC AGTTAGACGA   | 1980 |
|    | AGATAATTTA AATGAAATGA TGGTGACGAT CGAGTGAATG ATAATTTTAA AAAGCAACCG  | 2040 |
| 25 | CATCATTTAA TATATGAAGA GTTATTACAA CAAGGTATTA CTCTAGGTAT TACAACCTAGA | 2100 |
|    | GGAGATGGTT TAAGTGACTA TCCTAAAAAT GCTTTTAATA TGGCGAGATA TATTGATGAT  | 2160 |
|    | CGCCCATATA ATATTACTCA ACATCAATTG CAATTAGCTG AAGAAATTCT GTTTGATAGA  | 2220 |
| 30 | AAAAATTGGG TGTTTCCCAT TCAAAACACAT GAAAAATAAG TCGCTTGTAT TACAAAGGAT | 2280 |
|    | GATATAGGCA CAAATATAGA CACTTTAACT GATGCGCTTC ATGGTATTGA TGCGATGTAC  | 2340 |
|    | ACATATGATA GTAATGTCTT ATTAAOGATG TGTATTGCAG ACTGTGTACC AGTATATTTT  | 2400 |
| 35 | TATAGTACAA AACATCATTT TATTGCATTG GCGCATGCAG GTTGGCGTGG TACCTATACT  | 2460 |
|    | GAAATTGTAA AAGAAGTGCT AAAACATGTG AACTTTGATT TGAAGACTT ACATGTGCTT   | 2520 |
| 40 | ATTGACCATT CTACATCATC AAGTTATGAA ATTAATGATG ATATTAAAA TAAATTTGAA   | 2580 |
|    | ACATTGCCAA TTGATAGTGC CAACTATATT GAAACTAGAG GACGAGATCG TCATGGTATT  | 2640 |
|    | GATTTGAAAA AAGCCAATGC TGCAATTATTA ATTTATTATG GTGTTCTTAA AGAAAATATT | 2700 |
| 45 | TATACGACAG CGTATGCTAC ATCTGAACAT TTAGAATTAT TTTTCTCTTA TCGATTAGAA  | 2760 |
|    | AAAGGTCAAA CAGGACGCAT GTTAGCATTC ATTGGTCAAC AGTAAACAAG GAGGAGATAT  | 2820 |
|    | GTTCGCTGTG GAAAGATAAT TTACAACAAA TCTCAACACA AATTAATGAC AAAAGTGAA   | 2880 |
| 50 | AAAAATATTT TTCAACAAAA CCAACGTGA TTGCAGTTAC AAAATATGTT ACAATAGAGC   | 2940 |
|    | GAGCTAAAGA AGCGTATGAG GCTGGAATAA GACATTTTGG TGAGATAGAA TTGGAAGGCT  | 3000 |

EP 0 786 519 A2

AATCTCGAAA AGTTAAGGAC GTTATAAACG ACGTAGATTA TTTCCATGCT TTAGATCGAT 3120  
 TGAGCTTAGC CAAAGAAATT AACAAACGTG CAGAACATAA AATTAAATGT TTCTTGCAG 3180  
 5 TGAACGTTTC GGGAGAAGCT TCTAAACATG GTATTGCTTT AGAAGATGTT GATCAGTTTA 3240  
 TAGATGATCT TAAAAAATAT GACAAAATCG AAATTGTAGG TTTAATGACG ATGGCACCAT 3300  
 TGACAGATGA TGAAGCATAT ATTAGATCGT TATTTAAACA GTTACGTTTG AAAAAAGAAG 3360  
 10 AAATACAACG ACTCAATTTA GAATATGCGC CTTGTGATGA ATTATCAATG GGAATGAGTA 3420  
 ATGACTATCT TAITGCAGTT GAAGAAGGTG CGACGTTTGT TAGAATTGGG ACTAAACTTG 3480  
 TAGGAGAAGA GGAGTGAGCC ACTTGGCTTT AAAAGATTTA TTTAGTGGAT TTTTGTAAAT 3540  
 15 AGATGATGAA GAGGAAGTAG AAGTACCTGA CAAACAACAA CAGGTAATG AAGGCCAGC 3600  
 AAAAGAGCAG TCACAACAAA CAACAAAACA AACGCAATC AAATCAGTCC CTCAAAAATC 3660  
 20 TGCATCAAGA TATCAACAA CGTCAGAAGA AAGGAATAAC CGTATGTCTA ATTATTCAAA 3720  
 AAATAATTCA CGTAATGTTG TAACATGAA CAATGCTACA CCAACAATG CATCACAAGA 3780  
 AAGTTCAAAA ATGTGTTTAT TCGAACACAG TGTTTTTTCA GATACACAAG ATATTGCTGA 3840  
 25 TGAGCTTAAA AACCCCGGTG CGACACTTGT CAATTACAA CGTATTGATA AAGTATCAGC 3900  
 GAAAAGAATT ATTGATTTTT TAAGCGGTAC T 3931

(2) INFORMATION FOR SEQ ID NO: 343:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 3150 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 343:

AATTGTGGG GGACTCTTAG GTTTTGTGAT GCAAGAACA AGATTTTGTT TAACAGGTGG 60  
 40 CTTTCGAGAT ATGTATGTGC AAAAGAATAA TAAGATGTTT TATGCAATTAT TAATCGCTAT 120  
 TACTATTCAA AGTATAGGAT TATTGATTTT GACGGCAACA GATATTTTAC AAATTCCTGC 180  
 45 ACATAGTTTT CCAATATTGG GAACAATTAT AGGTTCTTTT ATTTTGGAA TTGGAATAGT 240  
 ATTGGCTGGA GGATGTGCAA CAGGTAAGTTG GTATCGCGCT GGTGAAGGGC TAATTGGTAG 300  
 TTGGATTGCA TTAGTATTAT ATGCTGTTAC TGCAGCAATC ACTAAAACAG GGATTTTAAA 360  
 50 GCCAGTAATG GATAAAATTA ATCAACCAAC GAATGTAATAT AGTGATATGT CTCAAAACAC 420  
 TGCATTCCTG TTTTGGGGAT TAGTCGTTAT ATTAACATA ATCACCATT TCTAGTTGT 480

|    |  |      |
|----|--|------|
|    | AGGTATTAGA TATTACCTTT TCGAAAAACG ATACCATCCA TTTATTGCAG CAATTGTAAT  | 600  |
|    | TGGACTTATC GCACCTCTAG CTTGGCCAAT GAGTGCATCA ACTGGAAGAA ATGACGGTTT  | 660  |
| 5  | AGGTATAACA ACGCCTTCAG CAAATTTAGT ACACCTTTTG ATTACAGGTG AAACATAAAT  | 720  |
|    | TATTGATTGG GGTGTCCTTT TAGTTCCTAGG AATTTTCATT GGTTCATATA TTGCAGCTAG | 780  |
|    | AGGATCAAGA GAATTTAAAT GGCATTGGCC AGACAAGATT ACAATACGAA ACAGTGCCAT  | 840  |
| 10 | TGGTGGCATA TGTATGGGAT TTGGTGCCTC AGTTGCTGGT GGTGTTCTTA TCGGTAACGG  | 900  |
|    | TTTGGTTGAA ACGGCAACGA TGACTTGGCA AGGATGGATT GCGCTAGCAT gCGATGATAG  | 960  |
| 15 | TTGGTGTATG GACAATGAGT CATTTTATCT TTGTTCTGCC AATGAAAAAA GTACACCAAC  | 1020 |
|    | AATCTGCAAA GGTTAAACAG CAAACGCAAA TAGTATAGAA GATTATTATG CAAATGATGT  | 1080 |
|    | TGATCAAATA AAAGTGATTG GAAAAGGAGA AATAATTATG ATACACGAAT TAGGTACAGT  | 1140 |
| 20 | AGGAATGGTA TGTCCATTTT CGTTAATTGA AGCGCAAAAG AAAATGGCAA CATTGCAATC  | 1200 |
|    | TGGAGATGAA TTAATAAATTG ATTTTGATTG CACGCAAGCG ACGGAAGCCA TTCCAAATTG | 1260 |
|    | GGCTGCAGAA AATGGTTATC CTGTAACAAA CTATGAACAA ATTGATAATG CTTCTAGGAC  | 1320 |
| 25 | AATTACAATT CAAAAAGTTT AACGTTATCA TTTTAACAAT AAAATAGATA TTAGATTCTA  | 1380 |
|    | TGGCTACTTC CGCTAAATTA AAAGTGAGTA AGTAGCTTT TTTTTTTTAT TTCATGAAAT   | 1440 |
| 30 | CATTTTATA TAGTGTGGCA CATTTTATTC CAAAAGATGT AATAAAACTT AACGCATTTT   | 1500 |
|    | TGCTTTTAT AAATGTCTAG ATTATTATGA AAAAAAGGGA GTGGTAAGTA TGAATCTTAA   | 1560 |
|    | CGATACGATA TTTATGTTTT TGTGTACATT ATTAGTTTGG TTAATGACAC CAGGATTAA   | 1620 |
| 35 | TTTATTTTAT GGTGGGTTAG TTCAATCTAA AAATGCGCTT AATACTGTCA TGCAAAAGTAT | 1680 |
|    | GGCAGCAATT GTGCTTGTTA CATTGTATG GATAACAGTT GGTTTTACAA TTAGTTTTGG   | 1740 |
|    | GAATGGGAAT TTATGGTTTC GAAATTGGGA ATATACTTTT CTTAATCATG TAGGTTTTGC  | 1800 |
| 40 | GACTCAAGAA GATATTAGCC CACATATTCC TTTOGCTTTG TTTATGTTAT TTCAATGAT   | 1860 |
|    | GTTTTGTACG ATTGCAATTT CTATTTTATC TGGTTCAATC GCTGAGAAAA TGAAGTTTAT  | 1920 |
|    | TCCTTATTTA TTATTCGTAG TAATATGGAC TGCTCTTGTA TACAGTCCAG TAGCACATTG  | 1980 |
| 45 | GGTTTGGGGC GGCCTGTGGA TTAACAACT CGGTGTATTA GATTTCGCTG GAGGTACGGT   | 2040 |
|    | TGTTTCATAT ACATCAGGTG TTTCTGTTT AGTATTAGCT ATTATGATTG GAAAAGGAAA   | 2100 |
| 50 | CAAACTTTCT GAATCAACAC CACATAATCT TATCATIACG TTGA1TGCGG GTATATTGCT  | 2160 |
|    | GTGGATTGCT TGTATGGAT TTAATGTAGG TAGTGCTTTT ACATTGATA ATATTGCGAT    | 2220 |
| 55 | GCTTGCAATT ACAAACTAGT TCATTTCAGC CAGTGCAGGT GCTATAGGTT GGTTAATTTT  | 2280 |



5  
 10  
 15  
 20  
 25

|   |      |
|---|------|
| ATTAGTTGTC ATTACTCTG CAGCAGGATA TGTAAACATAT CTTAGTGCAA CAATAATGGC   | 2400 |
| TTTAATAGGA GGTATCTGTT GTTATATTGT CATTAAATTAC ATCAAGGTAA AACTAAAAATA | 2460 |
| TCATGATGCA TTAGATGCAT TTGGTATTCA TGGTGTGGT GGTATTATTG GTGCTGTTTT    | 2520 |
| AACAGCAGTT TTCCAAGTA AAAAAGCCAA TCCTGACATT GAGAATGGCT TTATTTATAC    | 2580 |
| TGTTGACATA CATATTATAC TTGTACAAAT ATTATGTGTA ACAGCAGTTG TAAATTTTAG   | 2640 |
| TATAGTCATG ACGTTTATTA TTGCGAAAGT AATTAAATTA ATTACACCAT TATCTGTTAC   | 2700 |
| GGACAAGAA ACGAATATAG GATTAGACAA GATTGTTTcC GGTGAACATG CTTACTTTGA    | 2760 |
| AGGTGAGCTA AATAGATTCA ATAAACATAT TCGATATTAG AATATATTTA CATAGAATAT   | 2820 |
| TCATTGTCCT GACATTTAAC TAAAGGTTGA TGTGGGACA TTTTGTIATA CAAAAGTTTT    | 2880 |
| ATTTTGAAT CTTTTATGA AAGAAGCAGA AATATTATT AAAGCGTTA CACATATGCT       | 2940 |
| AAAATAAGGC TAAGTGTAC AAATAATGAT AGGTGAATAA GTATGAAAAA TATATCTGAT    | 3000 |
| ATTGCCAAAT TGCGAGCGT TTCAAAAAAGT ACAGTATCTA GATTTTAAAT TAATGGATCT   | 3060 |
| GTCAGTAAAA AAACAAGTGA AAAATTAAACA AGAATTATAG CAGAACATGA CTATCAACCG  | 3120 |
| AATCAATTTG CTCAAAGTTT AAGAGCGAGA                                    | 3180 |

(2) INFORMATION FOR SEQ ID NO: 344:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 3719 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 344:

35  
 40  
 45  
 50

|  |     |
|--|-----|
| GTTATAGTGA AATTGACTCA TCACATTTC AAGACCGTGA CAAACGCGTT ATTAGACGTG   | 60  |
| ATCATGTTAA AGAAGCACAA AGCTTAGTAG AGAACTATAA AGATACACAA AGTGCTGATG  | 120 |
| CTAGGATGAA AGCCAAACAA AAGTTAACA CATTAAAGCA ACCGCATCAA AACTATTTC A  | 180 |
| ATAAACAAAT TGATAAGGTT TATAATGGAT TACAACGCTA ATCCAAAGTA AATTATAAGT  | 240 |
| TATACATCTC GTTTTTAAAT GACAATTTAT CCCCGTAAAT ATTATAAATA ATCTTTTCAA  | 300 |
| ATTCCACATA GATATAGAGA CACTAATAAA CCTCTTTGTC TCGATATGAT AGTCTGCAAC  | 360 |
| GATTTCATGT GTAGGCTTTT TAAATTTTACA AATAAGGCTA AATATATAAG TTCTGGCACC | 420 |
| TAAATATAG AAAATACATA AAGTAAGTA TAGTTATTTT ATTATAATTA TAAATTTTTT    | 480 |
| ATTAATTAAT TGTAATAATG TATAATTATA ATTAATTAAC GTTTAATATT AAAATTAAC   | 540 |

|    |  |      |
|----|--|------|
|    | ATCGTTTCAA TATTACTTAT AGGGATGGCT ATCAGTAATG TTTCGAAAGG GCAATACGCA  | 660  |
|    | AAGAGGTTTT TCTATTTCGC TACTAGTTGT TTAGTGTTAA CTTTAGTTGT AGTTTCAAGT  | 720  |
| 5  | CTAAGTAGCT CAGCAAATGC ATCACAACA GATAATGGCG TAAATAGAAG TGGGTCTGAA   | 780  |
|    | GATCCAACAG TATATAGTGC AcTnCAACT AAAAAATTAC ATAAAGAACC TGGCAGATTA   | 840  |
| 10 | ATTAAGCGGA TTGATGGTGA TACGGTTAAA TTAATGTACA AAGGTCAACC AATGACATTC  | 900  |
|    | AGACTATTAT TGGTTGATAC ACCTGAAACA AAGCATCCTA AAAAGGTGT AGAGAAATAT   | 960  |
|    | GGTCCTGAAG CAAGTGCATT TACGAAAAAA ATGGTAGAAA ATGCAAAGAA AATTGAAGTC  | 1020 |
| 15 | GAGTTTGACA AAGGTCAAAG AACTGATAAA TATGGACGTG GCTTAGCGTA TATTATGCT   | 1080 |
|    | GATGGA AAAA TGGTAAACGA AGCTTTAGTT CGTCAAGGCT TGGCTAAAGT TGCTTATGTT | 1140 |
|    | TATAAACCTA ACAATACACA TGAACAACTT TTAAGAAAAA GTGAAGCACA AGCAAAAAAA  | 1200 |
| 20 | GAGAAATTAA ATATTGGAG CGAAGACAAC GCTGATTGAG GTCAATAATG CTCATTGTAA   | 1260 |
|    | AAGTGTCACT GCTGCTAGTG GCACCTTTAT AATTTT TAGA TCACGATATG ATTTATTATC | 1320 |
|    | AATTGCAAT TAAAAAGTA AATAGTATCA AAGTAAGTG TATTTAATAT TAGAAAAATA     | 1380 |
| 25 | AAATTTTAAA TTTAGTATTA AAATGGAATG TTACTATATA GTTCAATGTG TATTATCACA  | 1440 |
|    | GAAAAATAAA TAATGCTTTA CTCTATATT TAAAAGTGTA TAATGAAAGT TAAGTAATAA   | 1500 |
| 30 | AGAGCGTGAA GAAAAATGTG AGTTATTAT ATAGAATATT CTCCTTTTCA TTTATGAATT   | 1560 |
|    | TGTTACAAAA TATTTAGTGC AAAAGCAGCA CGGAGGTATT CAATATGAAT AACGGTACAG  | 1620 |
|    | TTAAATGGTT TAATGCAGAA AAAGGTTTTG GTTTCATCGA AAGAGAAGT GGTAGCGAGC   | 1680 |
| 35 | TATTCgTACA CTTCcAGCA ATCGCTGAAG ATGGATACAA ATCATTAGAA GAAGGCCAAA   | 1740 |
|    | AAGTTGAATT CGACATCGTT GAAGGCGACC GTGGCGAGCA AGCTGCAAAAC GTAGTTAAAA | 1800 |
|    | TGTAATTTTA ACTTATTCAA ACAGTCCTTA CTATAGGGCT GTTTTTTAT GCTTTAAATC   | 1860 |
| 40 | GATAACAGTT GGTGTGGTAA AAGCACTAGC CGTTATTTTT TTGTCCAATA AATTTAGTTG  | 1920 |
|    | GAGATTTTAAc AATATATAAT GGTCTAAAA TAAATCGAAC TGATGAAAA GTTTTTTACT   | 1980 |
|    | TTTCATCTGT CGACTTTTG ATTTTGAATA TAAAAAGCG CCAATACAGA ACTTTAATAA    | 2040 |
| 45 | TGACGAGAAT TAAAGTCTGT ATATGGCGAT AACAGAAGT AATGTTAAAC ACTCAAAATG   | 2100 |
|    | TTTAACAATA ATAGGATACC ACATCGCATA ATATCTTACT ACTTAATTAA TAATTTTAACT | 2160 |
| 50 | AATCAACTTT TTGTTAATTT TTTATTAAAG CTGATTAAAT ATTGAGAATA TTTATTGTTT  | 2220 |
|    | TTAAATCTC ATAATAATTC AGTAATCTTG TTTTCATTTA AAAGCGGAAA CATTTAAATA   | 2280 |
|    | ATTAATAATAA AATATTGCGT TTAATTTACA CGGTCAAATA TACTTATTTC TAATGCTTTG | 2340 |

ATTTCTGTGA GTATTGGAA GCTACCATTG GCGAACGGTT TAACAATAGA CAATTGCTTT 2460  
 TCCGCTTGTT GTATTA AAAA AGGTTTGTGTA GATTGATTAT TAATATGCCA TTCCTCATG 2520  
 5 TATGTTTTTC ACTCTGCTT TAAAAAGGG TTAGAAAGTT TATAGTTGAG ACATTGATGT 2580  
 TCAACCAAAA TTTTGTTGA ATTCAATAAA TGCTTGTTT AAAATAGAAA TATTGTAAAT 2640  
 GTTATCGTCC AAAACTTCAC CAGTTAAGTA TTTGTTTGA ATTAATAATT GGCAGTTAGT 2700  
 10 TAAGAAGTCT TGATAATCAC GATCGCAAAA ATAGTTTTCG CGTGCACTTT TAGCATCGCC 2760  
 AAAAAAGTTA GCGACTGTTT CTGTTTCTCC TTTATTGAA CGTTCAATAT ATAATTGTGA 2820  
 AAATTAGCT ATTGTACT TTTGTTCTTT AGTTAGTTCA TTCAAAATAT TGGGCTCCT 2880  
 15 GAAATATCAT TTGTAATCTA TACCAATTT ATTGCAAAAC AAAAATAAT TTAATATTT 2940  
 GATGAAACTG TGTTAATAAG CTTTAACAAG CCTTAGTTG TATGATCTA TAAATATTATC 3000  
 20 TTAAATTGCA TAGGGTAAA TAATATGTAG TCCATAACTT TTAAGTATT TTCACTTAC 3060  
 ACCAAATTTA TAAGCTTGGT AGATAATTTT AGTACAATAC GTAAATTTT TGCTGTTCAA 3120  
 ATTTAATGTA ACTAGATAAC GATGATTGTT ATTCTCATAG TTTTCTTAA CCCATTGAC 3180  
 25 CGCTTTTTTA CCGCACCAG GATAGCTGCA ACGATAAACT TTCACTCAAT CATTTTGGC 3240  
 ACTGTCATA TTATATTAA AAGATTGAA GGATTGTGTA GTTGGTTTGT CGCCAGGCC 3300  
 CTCAAATTTG AAAATCGTTT TATCATCAAT CGCGATACTA CAATGACCA AAAATGACA 3360  
 30 CATGACAGGG CCTTTGTAA CAATAATATC ACCAGTTGT AATTGGAATT TGTCATCTTG 3420  
 AATTTCTGAA TACTTATTAT CTGCAATTGT TTTGGTGAG TTTATTGGG ATACGACAAC 3480  
 GAATAATATA AGTAAATTA TCGTTCGTTT AATATAGTTC ACTTAAAGC TCCTTGTGTA 3540  
 35 AGAAATATAT GTAAATAGTC TTAAATTAGA ATTGTAATCT TTAATAAGCT TGCAAGACTA 3600  
 AAACATATCT TAAATATTAA AGTATGAGAG TGTGAAATGT CTATTAGAA TAAATAACAG 3660  
 40 TCTGAACATC CATTGAGACG TTCCAGACTG GATATAAAAT GAATTTTATT TATAGACA 3719

(2) INFORMATION FOR SEQ ID NO: 345:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1676 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 345:

TTGCGTTGCC GCACCAAGAT ATTGAATGCC TAGCGATTCC GAGTATGCAA ACTGAACGCA 60

|    |  |      |
|----|--|------|
|    | CTCTTTAATA CGCGTTATCG CTTTTTGTA ATCTGCATCA TGATAACAAA TCATAACGCC   | 180  |
|    | ATAGCCACCT GCTGTCCGAA GATCATCTCC CGGCTTAATT ACTAACGGGA ATTCCECAATT | 240  |
| 5  | CITTAATCTCG TTTTCGAATT GCTCAATTTT TACAACTTTT CTTTTTGTA AAAACTTCCC  | 300  |
|    | ATTTGTCCAT TCAGGTATTC TTGCTTTATT ATTTAAAGCA ACAATAACG TTTTATCTAA   | 360  |
|    | TGCATAATAT TGCTGATTCA AGATTGTTT ATCATGAATA TATTGAAAAA AAATCTTTTT   | 420  |
| 10 | ATTTTCCTTA TGTGCCAATT GTTTGATCAA GTTTTCGTAA GATTGCTGAT TGTAAATGT   | 480  |
|    | ATAAATGAG TTCCGGAAGT CCTTACCAAT AACTTGAAT AGCTGATGCA ATTTGTCTGT    | 540  |
| 15 | CGCACTAGCT TCGTGAACAA TAACAGGTAA TTGATTGCT ATTAATAACT CCCTACCAGT   | 600  |
|    | TAAAAAATTA GATTGATGTT CGTCCGTTT CAACCATGGA TTGCATATAT ACGAAGGTCT   | 660  |
|    | TGACGTATAG ACAACATCTT TGTATATAA ATCACTTAAC GTTAAGTTTC GTCATTACC    | 720  |
| 20 | ATTATTGTG ATTACTTCCC ATTCCTTTT AAATGCGCAT GCTCTTCAAT AATGCTTGA     | 780  |
|    | TAAACGTCTT GATTTGTAA TAACCTAAC CCCATCAACG CCATTATTTT AGCGCTTTA     | 840  |
|    | ATTAACTGCT CATCACCATG TACACTCGCA GCCGCTTCTC TAAATCTATG CGTATGCTCT  | 900  |
| 25 | ACTAAATTAC GTGATCCTAT TTTAATATGA GGATGTATTG TTGGCACAAC ATGACTTAGC  | 960  |
|    | TTCCCTGTAT CCGTAGAGCC ATAACCAAAA TCATCATCAA TAACTGCTTC ACCAACTTCT  | 1020 |
| 30 | TCAGCATATT TAGCAAAATA ATCATCTAAT TTCGGCGTTT TAATGAATTC ATTCACACCG  | 1080 |
|    | TTTTGAATTC GACCAAAATC ATAATCACAA CCACTCTGTA TCGCAGCTCC ACGTGCAGT   | 1140 |
|    | TGATTTACTT TTTCTGTTAA TATATCCAAT TCTTTACGCG TCATTGCTCT AGTATAAAAA  | 1200 |
| 35 | CGAGCATGTG TATAGTCTGG AATAATATTA GCTGCTTTCC CGCCATCTAA AATCACACCA  | 1260 |
|    | TGCACACGTT GATCTTTTTT AATATGTTGT CGTAGTTGTG CTACACCATT AAAATACTA   | 1320 |
|    | ATCATAGCGT CTAATGCATT TAACGCTTCA TCTGCATTTT CAGAGGCATG AGCACTTTTT  | 1380 |
| 40 | CGGTAAAATT TAACATCTAA AACATCCACT GCCAAGTAT CAATCGTTTT ATAAGTTTCA   | 1440 |
|    | TTTCCCGGAT GAATCATTTA GGCAATGTCT ATTTGATCAA TCACACCAGC CTGACATAA   | 1500 |
|    | GAAGCTTTAG CGCTACCATT TTCCCACTCT TCTTCAGCTG GACATCCAAG AACGACTACT  | 1560 |
| 45 | TTACACCAA TTTGGTCAAT CACTTGCTTC AAACCAATTG CACCAAGAAC ACTTGACGTT   | 1620 |
|    | CCAATGATAT TATGACCACA AGCATGACCC AATCCTGGCA AAGCATCGTA TTCTGC      | 1676 |

## (2) INFORMATION FOR SEQ ID NO: 346:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1294 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double

# EP 0 786 519 A2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 346:

|    |  |      |
|----|--|------|
| 5  | TACAGTAGGA ATCATAAAAC CTAATACAAC AAATACAAA CCATTTAAGG CATAACTAAA     | 60   |
|    | TGTGTTCCAA ATTGTATGTT AATTCATTG TAGTCCGTT TGTGCTCTAA TTAAACGGTC      | 120  |
| 10 | GCGTTCTAAA CCATGGATTA GACCTGCGAT TACAACCTGA ATGATACCTG AAGCATGAAC    | 180  |
|    | TTCTTCTGCT AAAAAGTATA CGACAAAAGG AGTTAATAAT TGAATAAAAG TTAAGGTATT    | 240  |
|    | GTTATCTTTT AAACCTTTAT TAGCGGTTAA GTCTATACGT ATTCTAACGA CAACGAATCC    | 300  |
| 15 | AATAATTGCA CCAATAAGTA CACCTAGTAT TGTTGAAATG ATAAATTGTT CAACAGCTTG    | 360  |
|    | GAATAATGAA AAGGTACCAG TTAATAATGC AGTAACAGCA ATTTTAAATG AAATGATACC    | 420  |
|    | TGCTGCATCA TTGAATAAAG ATTCACTTTC TAAAAATGTC ATAGAACCTT TAGGTAATAA    | 480  |
| 20 | TTTTCCGCGT GTAATAGCAG ATACTGCTAC TGCAATCAGTA GGACATAAAA TTGCTGCTAT   | 540  |
|    | TGCCAAAAGCG GCTGGCATTG GTAAGGCAGG CCAAATCCAA TGTATAAAAT AGCCAACACC   | 600  |
|    | GACTACAGTT GCAAAACACTA GTGCCATTGA CATTAAATAGT ATAGGTTTAC GATAATTCTAA | 660  |
| 25 | TAATTTTGTT CGAGAGACGT GGGTACCTTC CACAAAAAGT AGTGGCGCGA TAACGGCAAA    | 720  |
|    | CATAAACTACT TCAGAATTGA ATTGGAATAT AACTTGTATT GGAATAATGA AAATAACGAC   | 780  |
| 30 | ACCTAATGCA ATTTGAATAA AGGCAGTAGG AATTTGTGGG AATCGATTAT TGATAACCGA    | 840  |
|    | ACTAATAATC ACAGCAAAAA TAAAAATTAA AAATGCTTCT AATAGTGCCA TACAATACTC    | 900  |
|    | CTCAAAATTT TAATAGTTAA TATTTTATCA CTTTAAAGGC ATAATGACAT AGATATATTG    | 960  |
| 35 | ATAAAATGAA GTTATTTTCA AAAAACTCT AGTATCGGTT GAACTGATAC TAGAGCGAGA     | 1020 |
|    | TGTTTAAATT ATTGATTGTC ATATCTGAAA TGACCGCTGT CATTTTGTGC TTGTTTCATAC   | 1080 |
|    | GCGAGCTTTT CAGCATTGCT TTTGTATTTT TTATAAAGa AAAATAaAAA TATnaAccaG     | 1140 |
| 40 | AATGGCGAAA TATAAATAGC TGCTCTTGtT TCGTCACTAA AGAATAATAA AATGAATACA    | 1200 |
|    | AAGAAGAAGA ACGCTAGAAT AATGTAAGCA ATAGGCTTAC CACCAATCAA CTTAAATTTA    | 1260 |
| 45 | CTGTTTTTAT GTGCCTCAGG ATGCTTTTTT AAAT                                | 1294 |

(2) INFORMATION FOR SEQ ID NO: 347:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1935 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

|    |             |            |             |            |             |             |      |
|----|-------------|------------|-------------|------------|-------------|-------------|------|
|    | ACATGATAAT  | GATGACGCTA | TTAAAAACAG  | TTTTTTATTT | TTCATTGTTA  | TAACCTTCTT  | 60   |
|    | TCGTATGATT  | GATATTTGTT | GATATGTATC  | GACATGTGAA | TAATATCACA  | AAAACAGAGA  | 120  |
| 5  | ATATATATTT  | AACATATTAT | TAAATGATTT  | TGTTAATATT | ATTAATAACT  | TTATCCTCTT  | 180  |
|    | TAAAAATAAT  | GTGTGTACAA | AGTCATTAAAT | TTAGCAAATA | TTTTTAITTA  | GTAGTTAATA  | 240  |
|    | ACCATCGATT  | TGAAATTTAT | ATATAATTAT  | TAGCTAAATA | ATATCCTGCA  | TCTTTCTCAT  | 300  |
| 10 | ACAAATTTACT | ATAAAATagC | ATATCCGATA  | TCAGCGTTAA | TAAGATCGTT  | GATACTAGmC  | 360  |
|    | AGTTAAATTC  | ATAGAACGAA | ATCAAAATAC  | ACACTACTTT | CTGCATTTTA  | AATTATGTTT  | 420  |
| 15 | AAGAATCAnA  | ATTATGTTTA | nATAAAATATA | TATACTACTT | TGAAAGGTGT  | GAGCTTAATG  | 480  |
|    | ACAACTTTTA  | GTGAAAAGA  | AAAAATTCAA  | TTACTAGCAG | ATATTGTTGA  | ACTACAAACT  | 540  |
|    | GAAAAATAAT  | ATGAAATAGA | CGTTTGTAAT  | TATTTAACAG | ATTTATTCGA  | CAAGTACGAT  | 600  |
| 20 | ATTAAATCTG  | AAATTTTGAA | AGTTAATGAA  | CACCGCGCCA | ATATGTTGCG  | AGAAATCGGT  | 660  |
|    | AACGGCTCAC  | CTATACTCGC | ATTGAGTGGT  | CATATGGATG | TTGTTGATGC  | AGGAAATCAA  | 720  |
|    | GATAATTGGT  | CATATCCCCC | TTTTCAACTG  | ACAGAAAAAG | ATGGCAAATT  | ATATGGCCGA  | 780  |
| 25 | GGCACTACAG  | ATATGAAAGG | CGGTTTAAATG | GCTTTGTCG  | TATCTCTAAT  | CGAATTAAAA  | 840  |
|    | GAACAAAATG  | AATTGCCTCA | TGGAACGATT  | AGATTACTGG | CTACTGCTGG  | CGAAGAGAAA  | 900  |
| 30 | GAACAAGAAG  | GTGCCAAATT | ATTAGCTGAT  | AAAGGCTATT | TAGACGATGT  | CGATGGCTTA  | 960  |
|    | ATTATTGCTG  | AACCAACTGG | ATCTGGAATT  | TATTATGCAC | ATAAGGGGTC  | TATGTCATGT  | 1020 |
|    | AAAGTAACCTG | CAACTGGTAA | AGCTGTCCAT  | AGCTCAGTTC | CATTATTATGG | TGACAATGCA  | 1080 |
| 35 | ATTGATACAC  | TGCTTGAATT | TTATAATCTA  | TTTAAAGAAA | AATATTTCAGA | GCTTAAACAA  | 1140 |
|    | CAAGATACTA  | AACATGAATT | AGATGTTGCG  | CCTATGTTCA | AATCATTGAT  | TGGAAAAGAA  | 1200 |
|    | ATTTCTGAAG  | AGGATGCAAA | TTATGCATCT  | GGTCTTACAG | CTGTATGTTT  | GATTATAAAT  | 1260 |
| 40 | GGCGGcAAAC  | AATTTAACTC | TGTACCAGAT  | GAAGCTTCAC | TTGAATTTAA  | CGTAAGACCA  | 1320 |
|    | GTTCTGTAGT  | ATGATAACGA | CTTTATAGAA  | TCGTTTTTCC | AAAATATCAT  | TAATGATGTG  | 1380 |
|    | GATAGCAATA  | AGCTTTCACT | CGATATTCCA  | AGCAATCACC | GACCTGTAAC  | AAGCGATAAA  | 1440 |
| 45 | AATAGCAAAAT | TAATTACTAC | GATTAAAGAT  | GTAGCTTCTA | GTTATGTAGA  | ACAAGACGAA  | 1500 |
|    | ATATTTGTGT  | CAGCGCTTGT | AGGCGCAACA  | GATGCCTCTA | GTTTCTTAGG  | AGATAATAAG  | 1560 |
| 50 | GACAAATGTT  | ATTTAGCCAT | TTTTGGACCA  | GGTAATCCAT | TAATGGCACA  | TCAAAATCGAT | 1620 |
|    | GAATATATTG  | AAAAAGATAT | GTATCTGAAA  | TATATTGATA | TTTTTAAAGA  | GGCTTCCATT  | 1680 |
| 55 | CAATATTTAA  | AAGAAAAATA | AGAAGCATGC  | TGTCAGCTGC | CCTATTGCGG  | TGCTGGCAGT  | 1740 |

TCAAAATATCA ACAAGCACAT TTTCATTGAT TAAGTGATGT AAAACTGAAA TTATTGTGCT 1860  
 GATTTGTCAT ACATATATTG ACTAATGGGC ATATAAAAAG ATAGCCTCTA ATAGThACAT 1920  
 5 AAACTCGTAA AAnCC 1935

(2) INFORMATION FOR SEQ ID NO: 348:

(i) SEQUENCE CHARACTERISTICS:  
 10 (A) LENGTH: 1351 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 348:

CCTTThCCTA AACAAITTTT AGATTTAGAC AACAAACGGA TTTTAATCCA TACATTAGAA 60  
 20 AnATTTATTT TAATTAATGA TTTTGAAAAA ATTATTATCG CGACGCCACn ACAATGGATG 120  
 ACGCATACGA AAGATACACT TAGAAAAATC AAAATTCTCG ATGAAAGAAT TGAAGTCATT 180  
 CAAGGTGGTA GCGATCGTAA CGATACAATT ATGAATATCG TTAAACATAT TGAATCAACA 240  
 25 AATGGTATTA ACGATGACGA TGTCAITGTG ACACATGATG CAGTTAGACC ATTTITTAACG 300  
 CATCGTATTA TTAAGAAAAA TATTCAGCT GCTTTAGAGT ACGGTGCGT AGATACAGTG 360  
 ATTGATGCTA TAGATACGAT TGTACATCT AAAGATAATC AAACGATTGA TGCAATTCCA 420  
 GtGCGTAATG AAATGTACCA AGGTCAAACA CCTCAATCGT TTAATATTaA TTTATTAAAA 480  
 GAAAGCTATG CACAGTTGAG TGATGAGCAA AAGAGTATTT TATCTGATGC TTGTAAAGATT 540  
 35 ATGTAGAAAA CAAACAAACC GGTTCGACTT GTAAAAGGTG AGTTATATAA CATTAAAGTA 600  
 ACAACACCTT ACGATTTAAA AGTAGCGAAT GCTATTATTC GAGGTGGTAT TGCCGATGAT 660  
 TAATCAAGTA TATCAATTAG TTGCACCTAG ACAATTGAA GTTACGTATA ACAACGTAGA 720  
 40 TATTTACAGT GACTATGTCA TTGTACGTCC TTTATATATG TCAATTGTG CTGCCGATCA 780  
 AAGATAITAT ACTGGTAGCC GTGATGAGAA TGTCTTATCT CAGAAATTGC CAATGCTTTT 840  
 AATTCATGAA GGTGTGTGTG AGGTCGTATT TGACAGTAAA GGTGTGTTA ATAAAGGTAC 900  
 45 AAAAGTAGTT ATGTTACCGA ATACGCCGAC AGAAAAAGAC GATGTCATTG CTGAAAACCTA 960  
 TTTAAAAATCG AGCTACTTCA GATCAAGTGG ACATGATGGG TTTATGCAAG ATTTTGTGTT 1020  
 50 GCTAAATCAT GATAGAGCTG TACCA'TACC TGATGATATT GATTTAAGTA TTATTTTCATA 1080  
 TACAGAGCTT GTAACAGTAA GTTTGCATGC TATTGTCGT TTTGAAAAGA AATCTATTTT 1140  
 AAATAAAAAA ACATTGTGTA TTTGGGTGA TGGTAACCTA ggTTACATTA CAGCCATTTT 1200  
 55

GAGTCACTTC TCATTGTG ATGATGTCTT CTTTATTAAT AAAATACCTG AAGGCTTAAC 1320  
 ATTTGATCAT GCATTGAGT GTGTGGGTGG T 1351

5

(2) INFORMATION FOR SEQ ID NO: 349:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 411 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

10

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 349:

TCATCAAGTC TACGATAAAT TAAGTCCATA TCTAAAGSCT CGGGGTCGAC AGTTTGTAAA 60  
 GTATAACCAA CTGCACAGTG GCTACAACGC ATATTACAAA GATTTGTAGT TGTAAATTGC 120  
 ATGTTACTTA AAGTTAATTG GCCATGTTCT TTAACATCGT TATATGCTTC CCATGGGTGC 180  
 TTTTGAATAC TTATTTTAGG CTGTGTATTA CGCATTTTAT AAACCTCTTA ATTGTTATTT 240  
 GATACCAATT TGATACCGTT TAATCAAATA TGCTCATAGC TTGATGTTTT TTATCAGTAT 300  
 ATAAATGAGA GTACGTTTGA ATTGTTTCTG TAATGTTAGA ATGCCTCATT AATTCCATTA 360  
 ATAAATACAT ATCTACACCA TTATTAATTA AaTAGCTAGC GTACGAGTGT G 411

30

(2) INFORMATION FOR SEQ ID NO: 350:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1639 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

35

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 350:

TCATTTTCAT AGGTATTATC GCAGATCAGC ATAATAATCC ATTCCATATG AGTcCTTATT 60  
 TTGGTTATGC AGCAGCTCTA TTGGCAACAA GTGGCATTGA CTATACGTAT GTAAGAATGG 120  
 CAATGTACAT GGATCCACTT AAACCATATT TACCAGAATT GATGnATATG CATAAACTGA 180  
 TTTATCnCG TGGCGATGGT CGTATTAAAT ATATTACTAG AAATGATATT GCTAGAGGTG 240  
 TCATTGCTAT TATTAATAAT CCAGATACTT GGGGCAACG CTACTTATTA TCAGGCTACA 300  
 GTTATGATAT GAAAGAAGCTT GCTGCAATTT TATCTGAGGC ATCaGGCACA GAAATTAAT 360  
 ATGAGCCCGT TTCAATTAGG ACATTTCGAG AAATGTATGA TGAACCTAAA GGCTTTGGTG 420  
 CATTATTGGC ATCAATGTAC GAOGCAGGAG CAAGAGGACT ATTAGACCAA GAaTCCAATk 480

55



TTAATAATAA AGGAGCGTTA TAGTGAATAT CATCTCAACA ATTCTAATCA TATTTGTGGC 600  
 ATTAGAGTTT TTCTATATTA TGTACCTTGA AACGATTGCT ACAACTTCCA AAAAGACTAG 660  
 5 CGAGACATTT AATATAAGCG TCGATAAAAT GAAAGACAAA AATATTAACC TACTTTTGAA 720  
 GAACCAAGGC GTATATAACG GTTTAATCGG AGTTTGTCTA ATATACGGTT TGTATTACAG 780  
 CAGTAATCCA AAAGAAATAT GCGCAGCTAT TTTAGTGTAT ATCATTGGCG TTGCTATTTA 840  
 10 TGGTGGCCTT TCAAGCAATA TTAGTATCCT TTTCAACAAA GGCACATTGC CAGTATTGGC 900  
 ACTCATATCA ATGCTTTGGT AAGTATTGGT GTTTGGGGGG GTGGAGATGT AGTCGGAGGT 960  
 15 TTGGAGGATT TGAGCGAATT GTGTGTGGAC TTTAGACTCA GAGTATTCCA TCCTAATTAT 1020  
 TTCAAGCAGA GGTGACAGTA GCGTTGCCTC TGTTTCCTTA TAAAAAAATT ATTTAATGAA 1080  
 GAAAACCCAT ATCTGATTTA ATTTTCAGCT GATAAATACT CCATATATTA GAATGGCTAC 1140  
 20 TTTATCTATT GCATCAATCC TTTAAACAAA AAAACCCATG ATTTGAAAT TCCCGTATGA 1200  
 TGGGGTTCCT ACTCTCATGG ATCAGTTAAA TAAATATTAT CACTATCAGT TTATTAATTC 1260  
 AATATTATTA ACAATATATG TAGTCGTAAG AGGAAAGAGG ACATGAGAAC TTCGGTGTG 1320  
 25 ATTGGCATT CATAACGCTT CCAAACATAT TATTTGGTAA CAATAAGAAA CTATTTACAC 1380  
 AATATATTTT GTATAGTAAA ATTATTTTAT AATATTTAAA TCCAATTGCA CAAGGAGTGA 1440  
 30 TTATCATGGT ACCAGAAGAA AAAGGTTCTA TTACTTTGTC AAAAGAAGCA GCTATCATAT 1500  
 TTGCAATCGC AAAATTCAAA CCATTTAAGA ACAGAATTAA AAATAACCCA CAAAAACAA 1560  
 ATCCATTTCT TAAATTACAT GAAAAcAAA AATCTTAATC ACTTTtATT ATAGcATTTC 1620  
 35 TAATCTCAGA AATGCTATA 1639

(2) INFORMATION FOR SEQ ID NO: 351:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1816 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 351:

AAAAATCGCAT ATAGTAATAT GAATAACCAG ATTGTATCTA CAAAAAGTA TATTGAAAA 60  
 50 CCAAGCGCAC CCATTAATAA TCGGAGAATA ATATAATTT TTCTATIAAA GTGATCGGTA 120  
 TCCGAAAATC TAGCAATAAT TGAATTTACT GTAAACTGGC TAATCGCTGC AGATGCTAGA 180  
 AGTAATCCAT ACTGATTGTG TGTCATACCT AAATCTTTAG TTGCAAAAG AACAAGATAT 240

TTCTYATATT GTAaTAACGC TGCAAAACATA TCCATAACCC CGCTTCTTAG AGCCCCTTTT 360  
 AATTnATnAA TTAGGGGCTC TTATGCAGTT GGTGCATTAG CAACCAACTG TATTCTTTTG 420  
 5 TCCCTTTTIA ATTTATTaAT TAGGGGCTCT TTTGCTGTTG GTGCATTAGC AACCAACTAC 480  
 GTTCAATTTA ACCGAATAGT TTAaAATTAA ATACAACCT TAAATTAGTC TAAaACTAGC 540  
 CCTTTGGTTG TTCAACAAG CTCGCCATGA GATTTACAAA AGAATCAACT TGTGGCAATT 600  
 10 GCAACATGCT CGGATCATAA CTCATAAATG TCGAACGAAT CAGCGTTTCA TTATCAATT 660  
 CTACTTTTTC AAATCaaAT TGTTCTTTGC TGATATTTTT CATCATAATT TCTGGCAAGA 720  
 15 TTGTAACACC TACCCACTA ATCAACATTT CTTTGCAAGT TGCTACTTGA TCCACTGTAA 780  
 TAGTTGCATG GTAATCTTGT TCTAAATTAT CGTTATACCA TTCTTTTATT TGATTATAT 840  
 AAATCGGATC AGCTTGAaAC TCTATAAATG GTAACCTTTG AACATCATCT CGTCTATTTT 900  
 20 TTGGAAAAAT AAAATAATGA TCATCATTAA ATAAATGTGT GTTAGCTAAA TTCATTACCT 960  
 TATTTCCAGC AGTTATCATA ACATGATAAT CTCATAGATT TGCTTTAATT GTTCAGTTG 1020  
 AACCAACTTG CACTTGTATT TCAACATTAG GAAATTGGGC ATTATATAGG CTCAAAACCT 1080  
 25 CAGGAAGTAA GGTTTGTCCA ATCAAGAAG AACACCCGAT TGATATTGTT CCATTCACCT 1140  
 CACCAATATG TGCCGTCAIT TTGTCAAAAA ATAATCGCTC TCTTTTCAAC ATGTCACGAG 1200  
 30 CATGCTCAAT AATCATTGTT CCTTCAGTTG TTGTAATCAA TTGTTTTTTT GTTCTGATAA 1260  
 AAATATCTAC TCCAAAAGCA TTTTCAATAG CTTTtagTCT TTGTGTAAAC GCAGGTTGAG 1320  
 ATATATATAA AATTTCAGCC GCTTACGTA ACGTTTTCGT TTCGTCTAAT GTTATTAGTA 1380  
 35 AACGATAGTC TTCAATCTTC ATAATTTCCT CCCATAAATT ATTCAATTAT TGAACCTTCA 1440  
 TGGCTACAAG CATTCTAGAG TTCATTACTA ACGAATAATT TCACCAATTT TATTGGTAGT 1500  
 GCTGCAGCTT GAATTACTTA GTTTTTCTTT TGTGTGTGTT GATTTTTAGT TTGATTATAT 1560  
 40 TGCTTAGGCT TTATTGTGTT GCTTTTTTCA ATATTAGTTT TATTTGTGG CTTTGTATGA 1620  
 TTTTTTGA GCTTTGCATT AATTTTATTA AAGCAGTACA TGATTTTCTT TTGGAATCCT 1680  
 45 TTAAATCAT TTTCTAACT TGCCATAATT TGATGTGCAA TCATATATGC TCCATGAaAT 1740  
 TGCTTTTTTG TAATTGCTC ACTTCTAAT GCAaACATTA AATCATCTTC ATCTACCAAC 1800  
 TCATaTcACC ACTGG 1816

50 (2) INFORMATION FOR SEQ ID NO: 352:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 9956 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 352:

|    |  |      |
|----|--|------|
| 5  | GGGTGTGCGT CGATGATTAA TAAAGTATGG GTATACCATT AAGAATAACG CTACCCAAAT  | 60   |
|    | gaktgctagt GACGTGCCGC CAATGACATC TGAAAAGTAA TGTGCATGAA AATAAGGCG   | 120  |
|    | ACAAAATAAT ATGCTAAGCC ATAATATTCC CATAACCAAC GCACTCAACA CTTTGTGTAT  | 180  |
| 10 | TGTCTTAGCA GCAAGTGAAA TAATAATGAT CATTAAAGCG AAATATAATA ATGTGCTGGC  | 240  |
|    | GTTGGAATGT CCACTCGGAA ATGAAAAGCC TGTATCAAGC GCTAAATGAT TATATGGTCT  | 300  |
|    | TGGACGTAAT ACAGTATCTT TAATTAAATT GTTCATGATG ACACCTGAAA CCAATATATG  | 360  |
| 15 | CACAAACCAA ACCGCTAAAT GCCTCTGTTT AATAAACAGT ATGATTGTGA CGATAATGGA  | 420  |
|    | AATCAACACG ACACCTTTGA CATCTCCAAT TTCGCACTA AACGTCATAT AGTAATTAAA   | 480  |
| 20 | CAAATTGTTA ACATACTGAC GTTGIGGCTC ACCGAAATAA TCTGTAAACC ATGTTAATGA  | 540  |
|    | TCCCATATCT ATATTTTAA GCCATTCTTG ATTTGTCAC TACACTGTAA ACATACCTAT    | 600  |
|    | AAATACAATC AGCGCGATTA AAAATAAAGG CACTGTCAIT TTCGGTGATG TTAATTTTTT  | 660  |
| 25 | ATCTATCATC TTACAATCTC CTCGTATCAT CATTTTCATT TTACAATAGT TATCCATAAT  | 720  |
|    | ATCAATGTGC CACAAATTTT ACTTTACCGA CAATATCAAA ATTATAAAGT TCATATTGTT  | 780  |
|    | ATGTATATTG CAAATAAAAC ATTTGTATAAT TGAAATAACA ATATTTTGCT ATTTTCAATT | 840  |
| 30 | TAGTACGATT TATATTTATT ATACAGAGGG GGTAAAGCGT ATCAATAGAG TTAATTTTGT  | 900  |
|    | CTATATAGCG TTAATCATTa CATTAGTTAG ATTCCTTACC CCTATTATC CATCATTTTC   | 960  |
|    | AAATTTGATT TACTGGATAT TTGTATTATA TTTTATTCCT ATTATACTAT GCGTTATCGG  | 1020 |
| 35 | TTTCAAGGCC GAAAACTTA TTGCAACAAT GGTCAATATA CCTAATTTTT TAGGAATACT   | 1080 |
|    | TTATCGATTA TATGCTACG TCACACATAT ACTCTTTATG TAAAAGGATA GTGGGCATGT   | 1140 |
| 40 | CTCGCAACAA ATATGCTTGC GTCGACTTGT CACCATTGCG AAATTTTATG ACATACGCTT  | 1200 |
|    | TTCACGGGCT ATATTGAGAC CCACGCATTG ATCCACGTAA TAAACACATC ATGTAATAGA  | 1260 |
|    | AAAACAGCAC ACCCAATAT ATGGCGTTGC GCTGTTTAAc CAAGCATACT TCTATAGCTT   | 1320 |
| 45 | TAATAAGCCA GCAGAAGCAT ACCTAACCTT CTTAAATATG CTTTCCAAA TTATCCTCAA   | 1380 |
|    | GTTTGAATAC GATAATACGT TCACCTGTAA CTGTACTTAA ATCACTATGG AAGCTCATCA  | 1440 |
|    | CTTTGATACC TGTAATTTTA AAAATGATAT CATTCAAATC TTGCTCACCC GATTCAAATA  | 1500 |
| 50 | ATTGAGAAGC TGTTGTTTA ATATTTAATA ATCCTTCATT CGTACTACAT ACACGATATT   | 1560 |
|    | CAGCTGGCGT TAAGATACCT TGTAACATAA TAATCACCAT ATCTCTTAAA ATGTCGTGATT | 1620 |

55

|    |            |             |            |            |             |            |      |
|----|------------|-------------|------------|------------|-------------|------------|------|
|    | TTTCAGCTTC | GAITTCACCT  | TTGTTCTTT  | TCATATCACT | AACTCCAATA  | ATATTAAAA  | 1740 |
|    | TGATTACTTC | ATCTTTGTAT  | CGTTATCAGC | ACATCAACTA | TTACATTAAAG | TTTATCATTT | 1800 |
| 5  | TTAGTATATT | TTAAGAAGCT  | AGAACATTGT | AGATATGATG | ATATAITTAGT | TACTTAGCAT | 1860 |
|    | CGCAACATAT | CATCGTTAAA  | TCCAACTTT  | AAAACGCCCT | TCCTCATTA   | CGCTCATTA  | 1920 |
|    | ACGCAGCCAA | TGATTAGACA  | CCTTCTAGC  | GAAATGCTCA | TTATTGCGCA  | GTAGTCTTGC | 1980 |
| 10 | TACAACATAG | TCGGGTGCGT  | GAATAACGAC | AAGTAAACGA | ATTGGCGAAT  | GATACATCGT | 2040 |
|    | CCGATCAGCA | GCCATAACAG  | ATTGCCATGA | TAAGCCATAC | ATCAGATCAC  | TCGCATTACC | 2100 |
| 15 | TTGCATGACA | CCAACACCTG  | ACGTGACGGT | TTGTGTCGCT | TTATTCCCAC  | TTCCGTAAAA | 2160 |
|    | ATGCGGCGCA | ACTGTCGACG  | CATAATATTG | TAAATTAATC | CATTGTGCCA  | CAAGTCCCGG | 2220 |
|    | ACCAGAAATG | ATGGTATTTA  | ATAATGTGCC | ATCTTTATCT | TTACGCCAAT  | CATAATTGTG | 2280 |
| 20 | TAAAAATGTC | CGCCCTTCTA  | AATCAATGCC | TTTTGTTAAT | TGGCGTCGTC  | CAATTATAAA | 2340 |
|    | TGATGCATTT | TTAGCCAATC  | CCCATTCTGG | ACGTACCTCA | TCCTCAATCAC | TCGCAAAACG | 2400 |
|    | CTGCGCTTCT | TCCACTGGAT  | GATTACACAG | ACCAATCGTT | GGCAGTTTGT  | CCAAACGTTT | 2460 |
| 25 | GCGAITCGcg | TGTCAGAAA   | TCATCGGCAT | CGCGTCATTC | AATGATTTCAT | ATGCATCTAA | 2520 |
|    | AGCAATAGAA | GATAATGTGT  | CTGGCACATA | TACCCATGCC | AACGTATCAG  | TAGACGTATG | 2580 |
| 30 | ATGTTCTGCT | ACCGCAAAAA  | CAGTTGTCTC | TGGAATATAC | ACACCTGATT  | GTTTTAATCC | 2640 |
|    | TTGTCTGACA | TTTGACGAT   | TACATATCAT | CGCTAATAAC | TTAGCATTA   | AACCGCTTGA | 2700 |
|    | TGGCCACCA  | CAAGCCCCAC  | ATTCAGTGA  | TGCATGATGT | GGATTATTGT  | GAGAATGACT | 2760 |
| 35 | AGCATGACCT | GCTAACACAA  | CGAACGCGC  | AAATGCTTCG | GTTAAATCCA  | TCAATTTCAA | 2820 |
|    | CGCTGTAAC  | GCGAAATCAA  | TTTGTCTTGT | CTCAGTAAAT | CCAACAGGTA  | AGTCTGATGT | 2880 |
|    | TCGGTCAAA  | TCACGATCAA  | TCGTCAACTT | TGTTTCAGGC | TTTTTCAACC  | ACTTTTGTTT | 2940 |
| 40 | TATTTTTTGT | AAAGACGCGC  | GACTTTTCT  | AGGCATAATC | GAATTGACAA  | TGGTACTTAA | 3000 |
|    | GCTTAAAAAT | GGCCCACTTA  | ATTACGGCAA | TAACAGACTA | GGCATGACAT  | TATTTTTCAT | 3060 |
|    | CAATTTAAAT | GTGTAAAAACA | TCGATGACAT | TGTCGTTTGC | TGTTGTGATG  | AAACATTCAT | 3120 |
| 45 | ATCGTAGCGG | TCTGCAAAAT  | CTTTAATGCG | ATATGCCGCG | GGTACCATGA  | CAGGTAATGA | 3180 |
|    | ATCATGTTTG | AATTGTTCTG  | CTACGGCATC | TTTTTGAATA | GGTAATCCAA  | AGAAGCCTGC | 3240 |
| 50 | AATACCAATC | GTTTCAAAGG  | GCCCTGCTGC | TTGATATGT  | CTACGAAATG  | GTTCTGAACG | 3300 |
|    | AACATCTATA | CAAAATGCAA  | TTTGGCTTT  | CGTTGATGTG | CCCATCTGAT  | TTAGCTCGCT | 3360 |
|    | ATTATTTTCA | TCAACTGCTT  | GTGTGTCATT | TAACAATACT | GAATGTGGCT  | GATTAGCGTT | 3420 |

|    |             |             |            |             |             |             |      |
|----|-------------|-------------|------------|-------------|-------------|-------------|------|
|    | TGCTTTAATT  | TTTTGTTTT   | ACTGAGATT  | GTATGTCATT  | TCCCAGGCAA  | TTAGCCATAA  | 3540 |
|    | ATTTTTAAAT  | ACATTTTAT   | TCATAGTTGC | TGCAAAATGA  | ATAAACGTTT  | GAATTTTCATT | 3600 |
| 5  | GACGTCATGT  | TGTAGTAATA  | CATCGCTAGG | CATATCACTG  | TAGTAACACC  | ATGATGCAAC  | 3660 |
|    | AGTTTGCTTA  | AACCAATTT   | CCGATCTACT | TTCAACAATC  | TTAGCGACTG  | ACTTAAACTC  | 3720 |
|    | ATCACCACCT  | AGCAATTGTT  | CGACAACATA | CCGAATTGCC  | AAATAATCCG  | TTAACCAATG  | 3780 |
| 10 | TTGTTCAAAG  | TGATGCTGTT  | GTGAACGGTA | ATACAACATA  | CCTGCCCAAC  | CCGGTAACGC  | 3840 |
|    | CAAAAGATGT  | CCITCAACAT  | AAGCTTGGTA | GTCTTCCTGA  | TCTATTGAAA  | AATGAGTTAA  | 3900 |
| 15 | TACTGACTCT  | ATCGTCAATT  | CAGGATCATT | GGTAAGCCT   | TTAATCACTT  | GGCGTGTGC   | 3960 |
|    | TTTAGTAAAA  | CTATGGTCAT  | GTTCGGCTAA | ATGCAACCAT  | GCATGGTAAA  | AACTTTGCTC  | 4020 |
|    | ACGCTTCGGC  | ATTGTCCAAC  | TCGATAGAAA | TTGATCGATA  | TAAAGTTTGG  | TCCATTTAAT  | 4080 |
| 20 | CATTTGACGA  | TTCACTTGTT  | CGCTAAGTGG | CTCACCTTGT  | TCATCTATTA  | TTGCATCACT  | 4140 |
|    | CATCGGACGT  | ACATCATAGT  | GATGATATGA | TTCAAGCCATA | TCACGTTTGG  | ATTTTTCATA  | 4200 |
|    | TAGTAGATCA  | GCAACCAACAT | CAACATTGTA | ATGATTTCATA | TATGATGCAG  | GTACGTCCTT  | 4260 |
| 25 | TAATGTTTTA  | ATGTTATCAA  | TATAAAGATT | GATGTAGTGT  | TGCGGGATAT  | TGTAGTGATG  | 4320 |
|    | TTCAAGTAAC  | ATATCAGTAA  | CAAGTTGATT | AAAGACACTT  | TCATCTAATT  | CACCACGTGC  | 4380 |
| 30 | CACAGCGCTT  | TCTATTAAATG | CTTTATTGCG | GAAAAATATCC | ACATCTCGAA  | CATCACGTAA  | 4440 |
|    | CCATTTTGGC  | ACATCTTCAA  | ACGTATCGCG | TTCTAATCCT  | TCCCATGGAT  | TTGCTGCTGC  | 4500 |
|    | AAAAATCGAA  | ATTGGTGATA  | ATGGTGTAAT | AACACGTTTC  | GCATTTTCAA  | TGACTGAATT  | 4560 |
| 35 | GATATTTAAT  | TGTGTTGTCA  | TACCTTTTAC | CTCCTATAAA  | TACTTCTTCA  | AATAATTCCG  | 4620 |
|    | ATGACTTTCT  | ATCGTTTTCG  | AGCGTGCTTC | ACCTAGATTA  | ACTAACCACA  | CGTACAATAC  | 4680 |
|    | CGCAAAAGCC  | TTAGAGTATC  | GATGCCGCGC | CACCCAATA   | CTTAATAAAC  | TGCCAAAGAT  | 4740 |
| 40 | TAAAAATAAC  | ACACTAATGA  | TGACACTCAC | TGTAGGCGCG  | GTGTGCGCAT  | GTGTTGTTAT  | 4800 |
|    | ATTTTGTAAAT | ACAGCGTAAA  | AATAATTATG | TGTGATGACG  | TAGATAAAATG | TCACGATTGC  | 4860 |
|    | AATCAAAATC  | ATACCAACAA  | GACGTGCCAT | GGTCCCTTTA  | CTAAAGGCTA  | CCATTGATT   | 4920 |
| 45 | CCAAGATACA  | AGTAATGACC  | ATCCTAGAAT | GAGTGCACCT  | AACACTTCAT  | ATGCACTTCT  | 4980 |
|    | GTCACACTCT  | ATCCAAAATA  | GAAATGCCAC | GATAATAGCT  | AATACACGTC  | CCATGACAAT  | 5040 |
| 50 | CCAGCCATAA  | CGCTCTTTAG  | CAGATGCTTG | TTTTGGAATA  | TTGAATCGCT  | TCACGATAGA  | 5100 |
|    | ACCTGATTGT  | AAAAATAATG  | TTGCTTTAAA | AATACCGTGC  | AATATTAAAT  | GAATAATCGC  | 5160 |
|    | TGCTGAATAT  | ACACCCAATG  | CACATTGAAC | TAAACATAAG  | CCCATTGTAC  | TCATCGTAGA  | 5220 |

|    |             |            |            |             |             |             |      |
|----|-------------|------------|------------|-------------|-------------|-------------|------|
|    | AGAAATACTA  | GAAAGGATAA | GTAATAATGA | TAACGCAAAAT | CCATTATCAA  | ATATCGGCGC  | 5340 |
|    | AAACGAGTT   | AGAATAACAC | CACCTGCATT | CACAATTCCT  | GCATGCATAA  | TTGCCGATAC  | 5400 |
| 5  | TGGCGTTGGT  | GCGTTACAG  | ATTCAATCAA | CCATCGATGA  | AAAGGAAATT  | GTGTCGCCGG  | 5460 |
|    | TATCATGACA  | GCTAATACAA | GTAGTACATT | CGTCAACAAT  | GACCATGTGG  | GATGAACCTAT | 5520 |
|    | ATGTTGGTGT  | ACCCGCCACT | CGCCAGTCGC | AATATAAATA  | GTTACAATTG  | CTCCAACGAA  | 5580 |
| 10 | TGCAAGCCAA  | CCACATAAAA | ATGTCATGCT | TGATAATTTT  | GCAGACTCAC  | GTGGCACTTT  | 5640 |
|    | CCAAAAACGA  | TAAACGTTCA | TCAGCAATGT | TAAACATAAT  | AATGTAATAC  | CCCAGCAGAG  | 5700 |
| 15 | TGCCATCAGT  | CTTAAGTCTT | CAGACATCCA | TGCTAAAGAT  | GCAAAACGAG  | TAATCGCAGT  | 5760 |
|    | GAACAATGGA  | AAGTAATGTC | TATAATGATG | ATCACCTAGT  | AAATATCGCA  | TTGAAAACTT  | 5820 |
|    | TTGAATAATA  | AAGCCAAGCG | CCATTACAAA | GCCAGCTAAT  | AACCAAGATA  | AACGATCTAT  | 5880 |
| 20 | TTTAAATGGA  | CCTAAGACAT | GTGACCATG  | AATACCGAAA  | AAGCCAATGA  | CTGCAAAATA  | 5940 |
|    | TACTGGCATG  | ACTAGTATGT | ATAAATGTAA | TTTAATATAT  | CTCATTTGGCA | TAACTGGTGC  | 6000 |
|    | TAAAAACAAC  | AAGCCACTTA | TCAATGCAAT | GATAAGCGCA  | ATAACAAACA  | GTGAAAATAG  | 6060 |
| 25 | CAATTGAAAA  | CTTAACACTG | CATAACCTCC | TTATTTCTAA  | TCTCTCGCAT  | AATTGCTTAT  | 6120 |
|    | GTATAAAAAAT | AAAAACCTAC | AATAGTAGAT | TCTGTACATA  | ATGGCAGAAA  | ATTACTATT   | 6180 |
| 30 | GCAGGTTTCA  | GTTTAACTAG | ACACTGCATC | ACGGTACGTT  | GATATACCTT  | GTTCGAGTGT  | 6240 |
|    | TCTCTTTAAG  | CGTGCTCCCA | TGCACATATG | TATATAAAAT  | GTTACTTCTG  | TCTGTTCAAT  | 6300 |
|    | TCATCTTCAT  | AAATATGCTT | TGCTAGACG  | AGACCTAACG  | TGTTATTTCGT | TTTAAACTTA  | 6360 |
| 35 | TAACATAAAA  | TATAATTAAA | TTTCTGCTTC | ATGTCAAATT  | CATGAGCTTA  | ACCTCTATTA  | 6420 |
|    | AACCAATGAT  | TGTAAGATT  | TTGTAAATGC | ACCTGTACAG  | TTAGGCAGTA  | TTTCCCGTCC  | 6480 |
|    | TTTTAAAAATA | AAAAATTGCG | AGTTATGATC | ATAACAATTC  | AAGTTAGGAA  | AAAAATCAAT  | 6540 |
| 40 | TACGCACAAG  | ATAACTATGT | ACAATGAAGT | TAACTCATAA  | GCAAGGAGG   | TAATCTTAAT  | 6600 |
|    | GGGTATCATC  | GCTGGCATCA | TAAAGTTAT  | CAAAAGCTTA  | ATCGAACAAAT | TCACTGTTAA  | 6660 |
| 45 | ATAAGATTTC  | ATAACAAACA | AAGGAGGTCT | TTCACATGGG  | TATCAATTGCA | GGAAATCATT  | 6720 |
|    | AATTCAATTA  | AGGATTAATT | GAGAAATTCA | CTGGTAAGTA  | AGTTATAAAA  | ATCTCATAGA  | 6780 |
|    | TATGAACATC  | TTATTTGAAG | GGGGCCATTC | ACATGGAATT  | CGTAGCAAAA  | TTATTCAAAT  | 6840 |
| 50 | TCTTTTAAAGA | TTTACTTGGT | AAATTTTTAG | GTAACAACTA  | ATCTCAAAACA | TTAACGATCA  | 6900 |
|    | ACAACATCAT  | ACTATGTAA  | ATCAACATAC | AGGAGGACAA  | AACGATGGCT  | ATTGTAGGTA  | 6960 |
| 55 | CTATCATATA  | AATCATCAAA | GCAATTATCG | ACATTTTCGC  | AAAATAAATT  | AAGCGAATTG  | 7020 |

|    |  |      |
|----|--|------|
|    | TTATTGATGT GAGGTGAGTC TTGTTAGITT GTTGCAAATA AATGGTCTTG GTGTTTTTTG  | 7140 |
|    | TATAGGACGT TCTTAGTGGG ACATACGGAA TATTCGTGAT CTTGTAGTC TGACGCGTTA   | 7200 |
| 5  | TATTTTTGTG GCGTGTTTTA TGTTTGATAC TCGAGTTCTG AGACATTCAT GATTTGGCAT  | 7260 |
|    | GCGAAATCTT AATGATTTTC ATGATCTAGC GCAAGATATA TTGGCCACGT GCGGAATTGC  | 7320 |
| 10 | GTGACAGTT TAGACTGAAA CACTCGTGTG ACCGTAAGTG TTAATAGTAC ATTGATAGCT   | 7380 |
|    | GCATTTACTT CACTCATTTT TATGACTGTT AAACAATGAT TGTACCTTCA ATTAACAGTT  | 7440 |
|    | GGTACGATGG TTTTGCCATT TTTTCATCAAC GTAAATATAA AAAGGACTAA GACACATACA | 7500 |
| 15 | TGTCCTAGCC CTATGGATAA AATGCAAAAT TCTGCTTTAT CAAAACATATC ACACCTTAGA | 7560 |
|    | TAGATTGAAA ACAAAAAGAT CCTAAGAACA CCTAACTTT TTAITTAATTG TCATAAAATTG | 7620 |
|    | CAACAATTA AGCCACAATT CAAAAATGAT TATACTTCAT TCAACTTATC GTGCTGGTCT   | 7680 |
| 20 | AATTTGCCAT TGATATGGAT CTTCAAATTG TTGCCAATCT GCATCAATTT CTTGCGCATT  | 7740 |
|    | GACTAAGCAT GCGTCGAGTT CTTTTGTAA TTTTCTTCA TCTAATCTG TACCAATAAT     | 7800 |
|    | GACAAATGT GTATAGCAGT CGCCATATTC TGGATCCCAT TCAGCTGCGA CATCTTGACG   | 7860 |
| 25 | TTCTGCTAAT ATTTGTGTTT GTTGCGCTTC AGACATACTA GCCACCCAAT ATGTAACTGG  | 7920 |
|    | ATGAATATTG CAAGATGACC CTGCTTGAGA TAATAAACAT GCTACGTGAT TGTATTGTGC  | 7980 |
| 30 | TAGCCATACG ATACCTTTTG ATCGAACGAC ATTATTGGC ATGCTTCTA ACCAATCATT    | 8040 |
|    | GAACCTTTTA GCATGGAAG GTAGACGACG TTTATATACA AACGATGATA TACCATATTC   | 8100 |
|    | TTCTGTTTCA GGTGTATGCG ATGCATGCCC ACCAGACTCA AGTTCITTGA TCCATCCTGC  | 8160 |
| 35 | TGACTCGCTC GCTTTTTCAA AATCAAAACG CTGCGTATTC AAGACTTCTT TTAATCTAC   | 8220 |
|    | TTCAGAATTT GTTGCTCTAA TAATTTTAGC AGTCGGTTGC AATGCGCTTA ACATTTTTTC  | 8280 |
|    | TAACCTCGCT AGTTCITCTT CACTAATTAA ATCAATTTTA TTAATAATCA ATACATCACA  | 8340 |
| 40 | AAATCAACT TGGTCAATTA ATAAATCAGC AATCGAACGC TCATCTGTTT CGTCAACGCT   | 8400 |
|    | TTGATCAGTA TCCATCAATA AATCTTCTGA GTTGATGTCA TGTACGAAGC GGTTAGCATC  | 8460 |
| 45 | CACAACTGTA ACCATTGTAT CTAACGGCA AATCGCTGTA AGATCAATGC CAAGTTCATC   | 8520 |
|    | ATCAATATAT GAGAAAGTTT GTGCAACAGG TACTGGCTCT GAAATCCCTG TTGACTCAAT  | 8580 |
|    | AACAATTTGA TCGATGCCAC CTTTTTTCAC TAAACGCTCA ACTTCITTTA ATAAATCGTC  | 8640 |
| 50 | TCTAAGTGTA CAACAGATAC AACCATTAGA AAGTTCGACT AATTTTTTCAT CTGTACGCGA | 8700 |
|    | TAGTCCCCCA CCATCTGCGA CAAGATCTTT ATCGATATTT ACTTCACTCA TATCATTTAC  | 8760 |
| 55 | AATTACCGCG ATACGTGAC CTTCTCGATT TTGTAAAATA TGATTTAACA ACGTTGTCTT   | 8820 |

ACTTCAATTT ATTGTAAAT AGGAATAATT CTGTTTTACA TTATATAGGA GCGTTTCCTC 8940  
 TTTCCGAATC TTCGATAATA AAAAAATAGT ATACTTAATT AAATTATTGA GCGCTTTACT 9000  
 5 TTATAATGGA GACAAAGATA TATCTCACGA AAGAGAATCG AGGTGTATAA ACATGTTATT 9060  
 TGTCAITTTA GTTTTATATG TTA CTGGTAT TGCATTITATT CTACTCAGTG TTTTGGGTTT 9120  
 10 AAAGACTGAA GGATATCTA CGAAACATAC TTTATATACC ATTGGCAGTG CTATTATAAC 9180  
 GATTGCTTATT TTCATTTCAA TTGGCTATGC CATTCAATAC TTAAGTCAG CGCTTTATGG 9240  
 TTTGTAAGGT GAAGGTGATG AGTAACGGGT AGTTCGGGAG AGGTAACTT GCGTTGATTT 9300  
 15 TGATAAAGTG ATCATAGCTT TTAGTACTTG AGGATTTTTA TTGTTGCTGT TACGAATGTG 9360  
 GTCATGTTTA ATGCGGGACA GTAATTTAAG TTGTTTTTTT ACAATTGAGA GTGTGATATT 9420  
 TCGATTCGGT TCGAATTACT TTACATGGGA ATAATATAAA TTA AAAAGAA GCGGCCTAGT 9480  
 20 GTCAGTTGTG AATATACTGA ACATTGGTCG CTTTATTAG TAGTATGATA TGTAGTTTAG 9540  
 CTATTAATTT TTTTCAGGTC ATCCTTAATG CTGCTATCT CAGACATGGC ACTTTTAACC 9600  
 CAATCTCCTT GAGCTGCACC TTAAAAATTA GCTTTAAAAA ctTCGCAATG TTGCGCCATT 9660  
 25 TGTTCATTA ATACTTTTTC TTCACCTTTT AATCCGTTTT CAATATCTTT GTATTATATG 9720  
 TTAGTTTCAG GTGCAATAAC TGTGCGAATA TTTCTTTTT GCGCTTCCAT TTAGATATG 9780  
 30 AGATTAAGTG TTCTACTGT AGTACTTATA TCTGGCATT TTAAGTTCAT ATCTGGTTCT 9840  
 ATTAGAGTCA TTAATCTCC TC AAAATTAT CAGTCACCTA GCTTATCTAA CTGCTTTTCA 9900  
 TAAGACTTTT TTAAGTCTTC TTATATTCT TCTAATTTC CATTCTTGCT TTCTGA 9956

(2) INFORMATION FOR SEQ ID NO: 353:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 2411 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 353:

TTTCTTTTAA CAGATTATC cCATTAATTG TTCTACAAGC CCAGGATGAG CAATATCATT 60  
 TTTAGCAACC ATTAATAAAC CAGAAGTATC CATATCTATA CGGTGAACAA TACCTGGAGC 120  
 50 AATTCTCCA TTAATACCTG ACAAATTTTT AATTGATAC ATTAACCAT TAACTAATGT 180  
 ATGGGTATAA TGCCCTGGTG ATGGATGAAC TACCATGCCT TTCGGTTTAT ATACAACCTG 240  
 AACATCGTCA TCTTCATAAT AAATATCTAA ATTTAAATTT TCAGGTAGAA TATCAGCTTC 300



|    |  |      |
|----|--|------|
|    | AACAACCTTTA TCGTTTGCAA CGACTAAACC TGCCTTAATC CAATCTTGTA TCTGTTACG  | 420  |
|    | AGACCAATCA TTATTTAATT CAGGCAGCAA CTTATCTACA CGCATACCTG TTTGTTCTTT  | 480  |
| 5  | ATCTGTAATG TTAATTCAT AAGTCTCCAT TACTTAACTT CCTTCTCCTT TTTATTGGAA   | 540  |
|    | GTATCCTTTA ATTAGGCAAT AATAATTAAT ATTACACCAA TTGTAAACT TGAATCTGCG   | 600  |
|    | ATATTAATAA TTGGAATATC ATAACCAAAA ATATTGTAT CAATAAGTC AACAACTTCT    | 660  |
| 10 | CCTGTAAAAA TTCTATCAAT AAGTTTCCA AGTGACCTG CAAAAAGTAA ACTAATAGCA    | 720  |
|    | ACTTGCATAA ACAAATTATA TTGAGCATCT TTAATAAAGA AATATACTAA GGCTATTAAT  | 780  |
|    | ATAATAATGG TAATAATAAA GAAAAATGTC ATTTTCCAC TCAATATTCC CCATGACGCA   | 840  |
|    | CCATTATTTT GATGTGATG TATGTTTAAA AAGTGGGTA TCACCTTCAA TGAATCTCCA    | 900  |
|    | ATTTTCTATT TAGTAGCTAT AATATATTTA GTAACCTGGT CAAATATAAC GACAAATACT  | 960  |
| 20 | GCTATTAAAA TGGAAGTGGC AATAAAATAT TTTTGTGCA TTTTCGTCC TCCAATCAAT    | 1020 |
|    | CGTCCATGAG ACAACTCTTT ATATTATAGC TTACACCTGC TAATAAAAAA AGTAAGCATA  | 1080 |
|    | TTACATTAAA TCTAATGTTA CTAACCTCAAT ACTTGATAAA CTACTATGTT TTGACATTAA | 1140 |
| 25 | ATATGAACCT AATTATTCAT TTATCATATT TAAGATGACA TTAAAAATTA GGAAAGCAGG  | 1200 |
|    | CTGGAACATA AATCCCTAAA AAGACAGTAG TAAGATATT TCTAATTAAA AATTATCTTA   | 1260 |
|    | CTGTGTTCTT CTATTTATAC AATACTTCGT ATTGAATGGC TTCGTATGCT CCATCTGGCA  | 1320 |
|    | CATTACTGTA AAATCTTATA AATAGAATTT TTGATGATGG GTCCCTTCCT AGGGTGCCGT  | 1380 |
|    | CTCAGCCTCG GCTCTCGACT GGCACCTGCT CCTCAGGAGT CTCGCCATTA ATACTACGTA  | 1440 |
| 35 | TTAACATGTA ATTTTACTTT TAAATACTTT AAAAAATAA GACATGAATC GTCTACACTT   | 1500 |
|    | AATTGGACAA ATTCTATGAG AATAGATAAT GTTAATTTAA GAAAGTAGGC TATTTTGAGT  | 1560 |
|    | TtCACTCGAA TGTCAGTTCC AGGAATAAAT AAAGTTAAAC GAGAGCTAGG TTTTGTATTA  | 1620 |
| 40 | ATGGCAATTA ATATAAGGAA AATAGCAGCT CAACGAGCTG TACATTATAA AATACATATC  | 1680 |
|    | AAAAAAGCTG ATTTCTATCA AATAATTAAT AGAAATCAGC TTTTTTACAT TGCCTAAGAA  | 1740 |
|    | CTTAATGTCC CAAGCCCTAA AACTTGTGTG TATTTATTTG ATTTAGCAGC GATACGTTTA  | 1800 |
| 45 | TATCTTAAAT ACATAAATGC TAAAAGTATA AACCAAAATC GAATAAAATA AATTGCACGT  | 1860 |
|    | CTTGATCAAA CATTATAAAA TAATAACCCG AACACAAAAA TGAAGAATAC AAATATTAAAG | 1920 |
|    | TAGCCCATAT ATTTGCCACC TAATAGTTTG TACGTAGCAT TTTTATGTAG ATCTGGGTTT  | 1980 |
| 50 | TTACGACTAT AATTGATATA TGCAATGATA ATCAGACCCC ATACAACTAA AAATAACACT  | 2040 |
|    | GTAGAGATGG TAGTCACATA CGTAAATACT TTTGTGCGAT CTGGGAAAAA ATAGTTTAGT  | 2100 |
| 55 |  |      |

TTATTCGTCT TAGAAAAGTT CGGAGGTGCT TGTyGTTGAC TTGATAAACC GaAAAGCATA 2220  
 CGGCTATTTG AGAATATACC ACTGTTACAT GATGAAGCAG CAGCGGTTAA TACTACAAAA 2280  
 5 TTAATCAAGC CCGCAGCAAA CGGAATTCGG ATCAATGCGA ATrATTtTnC GAATGGACTG 2340  
 TTATCAGGAT CAACCTTGCTG CCAAGGGGTA ATAGACATGA TAACCGCTAA CGCCCCAAGC 2400  
 TThnATATTA A 2411

## (2) INFORMATION FOR SEQ ID NO: 354:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 605 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 354:

GngAATTATT TTTAATAATG AAAGGATTAC TThCATGGGT TTTTACTAGG AnTACCCAGA 60  
 AGGTCAAAT ATTTTTGATG CGCTAAGTCA ATATGAAGTT AAGCGACGCG GCGATATGGA 120  
 25 AGAGGATCCA TCATATAAAC AACTCATTTC TTATTGTTTA CTGAAAAATG AGCATGGCGA 180  
 GATATTAGTG TATGAACGAT TATCTGGCGG TGGAGAAGCT CGATTGCATG GACAATCTTC 240  
 AATAGGTGTA GCGCGTCATA TGAATGATGT TCCAGGAGCA GAATCTATTA ACGAAGTATT 300  
 GAGAGTTAAT GCACAGAGAG AATTAGAAGA AGAAGTAGGT TTAAGTGAGC AAGATTCAAC 360  
 AAATATGGAA TATATCGGTT TTATTAATGA CGATAATAAT GAAGTGGGCA AGGTACATAT 420  
 35 TGGTGTGTA TTTAAAATCA CTGTAAGTAC GAATGATGTA GAAGCTAAAG AAACAGATAC 480  
 TTTACGAATA AAATGGGTG AAAAAGGCAA CATAGAGTCA TATGATGATT TCGAAACGTG 540  
 GAGTGCATTA ATCCTTCAAG ATTTATAATC AAACGAGGTG ACATATATGT CAGATATTAT 600  
 40 TCCAG 605

## (2) INFORMATION FOR SEQ ID NO: 355:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 668 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 355:

TTTATTAGCA CATCCAACT ATTCATATGT TGGACAATTT TTAACGAAC TAGGATTTAA 60

ACAATTAGAC ACTGAACATT TAGCTGATTT AAATCCAGAG CGTATGATCA TTATGACAGA 180  
 TCATGCTAAA AAAGATTCCTG CTGAATTCAA GAAGTTACAA GAAGATGCAA CATGGAAAAA 240  
 5 GTTGAATGCA GTTAAAAATA ATCGCGTGA TATTGTTGAC CGTGAATTTT GGGCAAGATC 300  
 TCGTGGCTTA ATTTCTTCTG AAGAAAATGCG TAAAGAACTT GTTGAATTAT CAAAAAAGA 360  
 ACAAAAAGTAA GGTGGAAGTA AATGGCTATA AAAGAAATAA GTAGCCAATC TGCCATAGAT 420  
 10 CATAAAAGAA AAGACGCGAC AACACTCACG TATATAGTGA GTTTGTGCTT TCTTTTATT 480  
 TGTATATATT TAAATATGGC GATTGGTTCT TCGAAAATTA ATTTTAGCGA TATCATTCAC 540  
 15 TATGTTACTG GTCATACAGA TACGAAAGCA ACGTTTTTAT TGCATAATGT ACGTATGCCA 600  
 AGGATGATTG CAGGGTTATT TATTGGCGGT GCATTAGCGG TATCTGTTT GTTAATGCAA 660  
 GCAATGAC 668  
 20 (2) INFORMATION FOR SEQ ID NO: 356:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 787 base pairs  
 (B) TYPE: nucleic acid  
 25 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 356:  
 30 ATACAAAAAA ACATATCGAA AATAAGCTA AAAGAACTA TCAAGTTCCA TATTCAATTA 60  
 ATTTAAATGG TACATCTACA AACATTTTAT CGAATCTTTC ATTTTCAAAT AAACCTTGGA 120  
 35 CAAATTACAA AAATTTAACT AGTCAAATAA AATCAGTACT GAAGCATGAT AGAGGTATTA 180  
 GTGAACAAGA TTTAAATAT GCTAAGAAAG CTATTATAC TGTTTATTTT AAAAATGGTG 240  
 GTAAAGAAT CTACAGTTG AATTCAAAAA ATTACACAGC AAACCTTAGT CATGCGAAAG 300  
 40 ATGTTAAGAG AATTGAAATT ACTGTTAAAA CAGGAACTAA AGCGAAAGCA GACAGATATG 360  
 TACCATACAC AATTGCAGTA AATGGCACAT CAACACCAAT TTTATCAAAA CTTAAATTTT 420  
 CGAATAACA ATTAATTAGT TACAAATATT TAAATGACAA AGTGAAATCT GTATTAAAAA 480  
 45 GTGAAAGAGG CATCAGTGAT CTGACTTAA AATTTCGAA ACAAGCAAAA TATACAGTAT 540  
 ATTTCAAAAA TGAAGAGAAA CAAGTAGTGA ATTTAAATC AGACATCTTT ACACCTTAAT 600  
 50 TATTTAGTGC CAAAGATATT AAAAGATTG ATATTGATGT AAAACAATAC ACTAAATCAA 660  
 AAAAAATAA ATAAATCTAA TAATGTGAAA TTCCAGTAA CAATAATAA ATTTGAAAC 720  
 ATAGTTTCAA ATGAATTTGT GTTCTATAAT GCAAGCAAAA TTACmATTAA TGaTTTAAGT 780

## (2) INFORMATION FOR SEQ ID NO: 357:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 534 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 357:

```

AAAGTAAAAA TAAATCTCCC TTTTAACTT TCGTTTCTGC CATAGCCATT GCTCTTCTCG      60
TGATAGTTGC TACAATACTC TTTCTTTCAC GGTAAAAATG TTCAACTTGT TCTGCTAAAA      120
ATGCAGCTTC TTCTTCGACG TCAGTCATCA ACAATTCGCa AGCTAATGAT GCGTCATCTA      180
AAGSACCTAC AGCATTAACT CTAGGTCCAA TAATAAAACC AATTGTTTCT TCATCAATAT      240
TGTCAATTGTA TCCCGCTTCT TTTAGCAATG CTTTAAACAGA GGTGGGACAT TGATCATTTA      300
AGACTTTTAA TCCTGTtTC ACTAATGATC GATTTTCATC AGTTAAGGAT ACTAAATCCG      360
CAATGGTACC TATCGCAACT AATGCTTTAA AATAATCAGG TACATTTtCA ATCAATGCTT      420
GTGCTAATTT GIATGCAACA CCTGCCACC AC AATTGTTG GAACGGATAA TTAACGATG      480
GATGCATTGG ATGTACGATT GCATATGCTT CTGGTAATGT ACTACCAATT TCAT      534

```

## (2) INFORMATION FOR SEQ ID NO: 358:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 3621 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 358:

```

GGTGAGTCAA ATTAATGAA TCTAATAAGT CATAACTATC TATTTGTAAT GTGCAACGCT      60
TAACGCATAT AAAAAATGAA TGTGCTGATA ATGATTTACT CAAATTAAAA GGTGATTTTT      120
ATTCAATGAT GAATGAAAGT TGCCTTTTAA TTTTGGTAA AAGTTAATGC GTCAGTGAAT      180
TGTGTAAGTT TTTCAAAAAG TAAAAAGAAA TAATAAAGGT GAATTATTAG AATTCCAAxAA      240
ATAATTCATT ACATTCATAA AGCATTTTAC AAATGGTAAG AAAATGAGTG TTACAAATCT      300
AAATATTGCA AAAGAAGCTG ATTTAGTCAC AAAAAATGTC CTATGTAATA ATTGAGAAAA      360
GATGCATAT ATACGGTCTT CTACTATTTC AAATGTAAAA GTTGCTTATT TGGGTGGCTT      420
TTTGTTTTAT AAAAGTATAA AATTTTACTA TAATATATCT TGTAAGAGAAC AATGAAATGA      480

```

|    |            |            |            |             |             |            |      |
|----|------------|------------|------------|-------------|-------------|------------|------|
|    | ATAGCAAAC  | GTATTACTTT | GATACAAAA  | TGGTTGTAAT  | AAATATTAT   | CGATATGACG | 600  |
|    | ACTTGAATAT | GATAAAGTGA | CATATTTATG | TATATGACTA  | TTTCGCAAAA  | TGTAATCGAG | 660  |
| 5  | GTAGAATTC  | TTGACAATTC | TGTCAGTTTA | TAAGATGTTA  | TAAATATGTA  | GTGTATAAGG | 720  |
|    | AGGCAACAA  | GATGACTGAA | GAATTCATG  | AATCAATGAT  | TAACGATATT  | AAAGAAGGTG | 780  |
| 10 | ACAAAGTCAC | TGGCGAGGTA | CAACAAGTTG | AAGACAAGCA  | AGTTGTTGTT  | CATATCAACG | 840  |
|    | GTGGTAAATT | TAATGGGATT | ATTCCTATTA | GTCAACTATC  | TACGCATCAT  | ATTGATAGCC | 900  |
|    | CAAGTGAAGT | TGTAAAGAG  | GCGACGAAG  | TTGAAGCATA  | TGTCATAAA   | GTTGAGTTTG | 960  |
| 15 | ATGAAGAAAA | TGAAACTGGA | GCTTACATCT | TATCTAGAAG  | ACAACCTGAA  | ACTGAGAAGT | 1020 |
|    | CTTATAGTTA | TTTACAAGAA | AAATTAGATA | ATAATGAAAT  | CATOGAAGCG  | AAAGTAACAG | 1080 |
|    | AAGTAGTTAA | AGGTGGTTTG | GTTGTTGATG | TAGGACAAAG  | AGGTTTGTGT  | CCGGCTTCAC | 1140 |
| 20 | TAATTTCAAC | AGACTTCATT | GAGGATTTCT | CTGTGTTTGA  | TGGACAAACA  | ATTCTGATTA | 1200 |
|    | AAGTTGAAGA | ATTGGATCCT | GAAATAATA  | GAGTCATTTT  | AAGCCGTAAA  | GCAGTTGAAC | 1260 |
|    | AAGAAGAAAA | CGATGCTAAA | AAAGATCAAT | TATTACAATC  | TTTAAATGAA  | GGCGATGTTA | 1320 |
| 25 | TTGATGGTAA | ACTAGCGCGT | TTAACTCAAT | TTGGTGCAAT  | TATAGACATT  | GGCGGTGTG  | 1380 |
|    | ATGGTTTAGT | GCATGTATCT | GAACCTTCTC | ACGAACATGT  | TCAAACACCA  | GAAGAAGTAG | 1440 |
| 30 | TTTCAATTGG | TCAAGATGTT | AAAGTTAAAA | TTAAATCTAT  | TGATAGAGAT  | ACAGAACGTA | 1500 |
|    | TTTCATTATC | AATCAAAGAT | ACGTTACCAA | CACCTTTTGA  | AAATATTAAA  | GGTCAATTCC | 1560 |
|    | ACGAAAATGA | TGTCATTGAA | GGTGTCGTAG | TAAGATTGGC  | AACTTTGGT   | GCATTTGTTG | 1620 |
| 35 | AAATTGCACC | AGSTGTACAA | GGACTTGTAC | ATATTTCTGA  | AATTGCACAC  | AAACACATTG | 1680 |
|    | GTACGCCAGG | TGAAGTGTTA | GAACCTGGTC | AACAAGTAAA  | TGTTAAAAATA | TTAGGTATTG | 1740 |
|    | ATGAAGAGAA | TGAAAGAGTA | TCACTATCTA | TTAANGCAAC  | ATTACCAAAC  | GAAGATGTTG | 1800 |
| 40 | TTGAAAGTGA | TCCTTCTACG | ACTAAGGCGT | ACTTAGAAAA  | CGAAGAAGAA  | GATAATCCAA | 1860 |
|    | CAATTGGCGA | TATGATTGGT | GATAAACTTA | AAAATCTTAA  | ACTATAAATT  | AATATTAAAT | 1920 |
|    | AGTCAACTCC | ACATGTTTAT | GATTGcATGT | GGAGTATTTT  | TATGTAACAA  | AATATACTCG | 1980 |
| 45 | GAATGATAAC | GTGGgACAAA | TTTAACTAAG | TGTTTAAAAA  | GATaEAGTIT  | TAAGTGCTGa | 2040 |
|    | TTTTATCAT  | TACAGTAATA | AACTCATTTT | GAATACACAG  | TCTCATGTGA  | TATTATTAAA | 2100 |
| 50 | AAGATATaAG | AAAGAGAGGA | AGTTAGCTTA | TGACTAAACC  | TATAGTAGCT  | ATTGTAGGTA | 2160 |
|    | GGCCTAATGT | AGGTAATCT  | ACAATTTTTA | ATAGAATAGT  | TGGAGAACGT  | GTTTCGATTG | 2220 |
| 55 | TGGAAGACAC | GCCAGGTGTA | ACACGAGATC | GTATTTATTTC | TTCAGGTGAA  | TGGTTAACAC | 2280 |

AAATTAGAGC GCAGGCAGAA ATCGCCATAG ATGAAGCGGA TGTTATTATT TTTATGGTTA 2400  
 ACGTGGGTGA AGGATTGACA CAAGCGATG AAATGGTCGC TCAAAATTTA TACAAATCTA 2460  
 5 AAAAACCGGT CGTATTAGCG GTTAAACAAAG TAGATAATAT GGAAATGCGT ACAGACGTGT 2520  
 ATGATTTTCTA TTCATTAGGA TTTGGTGAAC CGTATCCGAT ATCAGGGTCA CATGTTTATG 2580  
 GTCTTGGTGA CTGTAGAT GCAGTTGTTT CTCATTITGG TGAAGAGGAA GAAGATCCCT 2640  
 10 ATGATGAAGA TACAATTCGA CTATCCATTA TTGGACGACC AAACGTAGGT AAATCAAGTT 2700  
 TAGTAAATGC TATTTTAGGT GAAGATCGCG TTATCGTTTC TAATGTTGCA GGGACAACGA 2760  
 15 GAGACGCTAT TGATACAGAG TATAGTTATG ATGGACAAGA TTATGTTTTA ATCGATACTG 2820  
 CTGGTATGCG TAAAAAAGGA AAAGTATATG AATCAACTGA GAAATATTCA GTATTAAAGAG 2880  
 CTTTAAAGC GATTGAACGT TCAATGTTG TTTTAGTGGT CATAGATGCA GAACAAGGCA 2940  
 20 TCATTGAACA AGATAAACGT GTTGCGAGAT ATGCACATGA ACAGGTAAA GCAGTCGTGA 3000  
 TTGTCGTAAA TAAATGGAT ACTGTGAAA AAGATAGTAA AACGATGAAG AAATTTGAAG 3060  
 ATGAAGTACG TAAAGAATTC CaATTTTATG ATTATGCACA AATTGCTTTT GTGTCGTCTA 3120  
 25 AAGAACGCAC AAGATTACGT ACATTATTCC CTTACATCAA TGAAGCAAGT GAAACACATA 3180  
 AAAACCGTGT TCAAAAGTCA ACTTTAAATG AAGTTGTAC TGATGCAATT TCCATGAAC 3240  
 CTACACCAAC AGACAAAGGT AGAGTTTGA ATGTCITTTA TGCAACACAA GTTGCTATAG 3300  
 30 AACCACGAC ATTTGTTGTA TTTGTTAATG ATGTAGAATT AATGcATTTT TCTTATAAAC 3360  
 GCTATTTAGA GAATCAAATC CGTGCCGCTT TTGGTTTGA AGKTACACCA ATTCATATTA 3420  
 TAGCTCGAAA GAGAAATTAA CGATTGGGGG ATAACAATGA CTAAAATTAC CGTTTTTGGT 3480  
 35 ATGGGAAGTT TTGGGACAGC CCTTGCCAAT GTTCTIGCAG AAAATGGACA TGATGTTTTG 3540  
 ATGTGGGGTA AAAATCAAGA TGCTGTTGAT GAATTAAATA CATGTCATAC AAATAAAAAG 3600  
 40 TATTTAAAAAT ACGCGAAATT A 3621

(2) INFORMATION FOR SEQ ID NO: 359:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 643 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 359:

CTTTCGAAA TTAGGATCnG nGCTATCTTG GCCCAATTA CCAAGGGAAC TanTGCACC 60

|    |   |     |
|----|---|-----|
|    | AATGCTCTT TCATCTCCAT GCCCTGTTGC TCATTATTAA TAACACGGTC TATTAAACACA   | 180 |
|    | ATGGCATTGT TTAATACGAT TCCAATTAA ATTAGCATAC CAATTAAACT TGGTACTGAT    | 240 |
| 5  | ATTGTTTCTC CTGTGATTAA TAGTGAATA ATTACCCGA TAACGTGAAA TGGTAAAGAG     | 300 |
|    | AATAAAATTG TAAATGGTGC TAGGCCACCT TTAATGTAA TAACATGGAT TAAATATACG    | 360 |
|    | ATAATGATTG CAGCTAACAT TGCAAAGGCT AATTGTGTCA TTGCATTGTT AATATCATCT   | 420 |
| 10 | GATGCACCAC CGATATTAAAC CTTTACATTA TTCGGTTTAT CCAAATTATT TATTTTAGAC  | 480 |
|    | ATCACITGTC GTGTGTGTGCC ACCCACATCT TTATTTGTGA CTTTAGCAGA TACCGTCGTT  | 540 |
| 15 | GCATAATCTC CTGTCTCTTG CGTCAATTTA CTGGGTGTCG TTGTTTAAAC TAACTAGCG    | 600 |
|    | ATATCTCCCA ATTTAATCGT ACCACCAGTC GGCTTTTCA AAG                      | 643 |
|    | (2) INFORMATION FOR SEQ ID NO: 360:                                 |     |
| 20 | (i) SEQUENCE CHARACTERISTICS:                                       |     |
|    | (A) LENGTH: 2524 base pairs   |     |
|    | (B) TYPE: nucleic acid  |     |
|    | (C) STRANDEDNESS: double  |     |
|    | (D) TOPOLOGY: linear  |     |
| 25 |   |     |
|    | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 360:                          |     |
|    | TTTCAGGCGY GCAAGGGCT TTTCTTTTG CTTTAAAGAT TATGATTAT CGTGCAAAAGT     | 60  |
| 30 | TAAGTGGTCG TATATAGTTT TAGTTTTAAA AAGGTAATTA AATAAAATAG TTTGCCGAGG   | 120 |
|    | GAGATGTCAA AATGATTAAA ATACCTAGAG GGACGCAGGA TATTTTACCT GAAGATTCAA   | 180 |
| 35 | AGAAATGGCG TTACATTGAA AATCAATTAG ATGAATTAAAT GACATTTTAT AATTATAAAG  | 240 |
|    | AAATAAGAAC ACCAATTTIT GAAAGTACAG ATCTTTTTCG AAGAGGTGTT GGTGATTCAA   | 300 |
|    | CCGATGTCGT ACAAAAAGAA ATGTATACAT TTAAGATAA AGGCGATAGA AGTATTACAT    | 360 |
| 40 | TAAGACCTGA GGAACAGCT CGAGTTGTGC GTTCATATAT TGAACATAAA ATGCAAGGTA    | 420 |
|    | ATCCAAACCA ACCAATTAAA CTTTATTACA ATGGACCGAT GTTTAGATAT GAACGTAAAGC  | 480 |
|    | AAAAAGGACG CTATCGTCAA TTATATCAAT TTGGTGTAGA AGCTATTGTT GCTGAAAATC   | 540 |
| 45 | CTAGCGTAGA TGCAGAAGTA TTAGCTATGG TTATGCATAT TTATCAATCA TTTGGATTAA   | 600 |
|    | AACATTIAAA GCTTGTTATT AATAGTGTAG GGGATATGGC GTCTCGAAAA GAATATAACG   | 660 |
| 50 | AAGCGTTAGT GAaACACTTT GAACCAAGTAA TTCATGAATT TTGTTTCAGAT TGTCAATCAC | 720 |
|    | GTTTGATATC AAATCCGATG CGAATTTTGG ATTGTAAAGT AGACCGTGAT AAAGAAGCGA   | 780 |
| 55 | TTAAGACTGC ACCTAGAATC ACTGATTCTT TAAATGAGGA ATCTAAGGCA TATTATGAAC   | 840 |

|    |  |      |
|----|--|------|
|    | GTGGATTGGA TTATTATACA CATACAGCAT TTGAATTAAT GATGGATAAC CCTAACTATG  | 960  |
|    | ATGGTGCAT TACAACGCTT TGTGGTGGTG GCCGTTATAA TGGTTTATTA GAATTGCTAG   | 1020 |
| 5  | ATGGTCCAAG TGAACAGGT ATTGGTTTGT CGCTAAGTAT AGAACGATTA TTGCTTGCAC   | 1080 |
|    | TTGAAGAAGA AGGTATCGAA TTAGATATTG AAGAAAACCT AGATTTATTC ATTGTTACAA  | 1140 |
|    | TGGGTGATCA AGCAGATCGA TATGCTGTGA AGCTATTAAA TCATTGAGA CATAATGGTA   | 1200 |
| 10 | TTAAAGCAGA TAAAGACTAT TTACAGCGTA AAATTAAAGG ACAAATGAAA CAAGCAGACC  | 1260 |
|    | GTTTAGGTGC CAAGTTTACA ATCGTTATTG GTGATCAAGA ATTAGAAAAT AATAAAATCG  | 1320 |
| 15 | ATGTTAAAAA TATGACAACT GGTGAATCTG AAACAATTGA ATTAGACGCA TTAGTCGAAT  | 1380 |
|    | ATTTTAAGAA GTAGAGAGGG CGTTAAAATA TGAGTAAGAG AACAACTTAT TGTGGATTAG  | 1440 |
|    | TTACTGAGGC ATTTTATAGGA CAAGAAATTA CATTAAAAGG ATGGGTTAAT AATCGTCGTG | 1500 |
| 20 | ACCTTGGTGG ATCGATTTC GTTGATTCAA GAGATAGAGA AGGAATTGTA CnAGTCGTGT   | 1560 |
|    | TTAATCTGCG ATTTTCAGAA GAGGCAATTGA AAATTGCTGA AACAGTACGT TCTGAATATG | 1620 |
|    | TTGTAGAAGT TCAAGGTACA GTTACGAAGC GTGACCCTGA AACAGTTAAT CCTAAAATTA  | 1680 |
| 25 | AAACTGGCCA AGTTGAAGTA CAAGTTACAA ATATTAAAGT GATTATAAAA TCTGAGACAC  | 1740 |
|    | CACCATTTTC TATAAATGAA GAAAATGTTA ACGTTGATGA AAATATTCGA TTAATAATACC | 1800 |
|    | GTTATTTAGA TTTAGTCTGT CAAGAGTTAG CGCAACATT TAAATGAGA CATCAAAATTA   | 1860 |
| 30 | CACGTTCTAT TCGTCAATAT TTGGATGATG AAGGGTTCTT TGACATCGAA ACACCAGTAC  | 1920 |
|    | TAACGAAGTC AACACCTGAG GGTGCACGTG ACTATTAGT ACCATCTCGT GTTCATGATG   | 1980 |
| 35 | GTGAATTTTA TGCATTACCA CAATCACAC AATTATTTAA GCAATTATTG ATGATTAGTG   | 2040 |
|    | GATTTGACAA ATACTACCAA ATCGTAAAT GCTTCCGTGA CGAAGATTTA CGTGACATC    | 2100 |
|    | GTCAACCTGA ATTTACACAA GTCGATATTG AAATGAGTTT TGTAGACCAA GAAGATGTGA  | 2160 |
| 40 | TGCAAAATGG TGAAGAAATG CTTAAAAAAG TTGTTAAAGA AGTTAAAGGC GTTGAAATTA  | 2220 |
|    | ATGGCGCTTT CCCACGCATG ACATATAAAG AAGCGATGCG TCGCTATGCT TCTGATAAAC  | 2280 |
|    | CAGATACACG TTTTGAAATG GAATTAATTG ACGTTTCTCA ATTAGGACGT GATATGGAAT  | 2340 |
| 45 | TTAAAGTATT TAAAGATACT GTTGAAAATG ATGGTGAAAAT TAAAGCAATT GTCGCTAAAG | 2400 |
|    | GTGCAGCTGA ACAATATACT CGTAAAGATA TGGGATGCTT TAACAGAAAT TGTAAACATC  | 2460 |
| 50 | ymTGGtGCTA AggtTAGCGT GGGGTTAAAG TTGTGGGAAG GTGGTTTTGA CAAGGTCCCA  | 2520 |
|    | ATGG   | 2524 |

(2) INFORMATION FOR SEQ ID NO: 361:



(A) LENGTH: 1507 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 361:

|    |  |      |
|----|--|------|
| 10 | TCGTTGAGTA AAGATCCAGA AAATTGGATG AGTAAACTTG ATGATGGAAA ACAITTTAACT | 60   |
|    | GAGATTAATA TACCGGGTTC ACATGATAGT GGCTCATTCA CTTTAAAGGA TCCAGTAAAA  | 120  |
|    | TCAGTTTGGG CAAAGACTCA AGATAAAGAT TACCTTACCC AAATGAAAGTC GGGAGTCAGG | 180  |
| 15 | TTTTTTGATA TTAGAGGTAG AGCAAGTGCT GATAATATGA TTTCAAGTTCA TCACGGCATG | 240  |
|    | GTTTATTTGC ATCATGAATT AGGAAAAATT CTCGATGATG CTAATATTA CTTGAGTGCT   | 300  |
|    | TATCCAAACG AAACAATTGT GATGTCATG AAAAAAGACT ACGATAGCGA TTCTAAAGTT   | 360  |
| 20 | ACGAAGACAT TTGAAGAAaT TTTTAGAGAA TATTATTATA ATAACCCGCA ATATCAGAAT  | 420  |
|    | CTTTTtACA CAGGAAGTAA TCGGAATCCT ACTTTAAAAA AAACGAAAGG TAAAATTGTC   | 480  |
|    | CTATTCAATA GAATGGGGG TACGTACATA AAAAGTGGTT ATGGTGCTGA CACGTCAGGT   | 540  |
| 25 | ATTCAATGGG CAGACAATGC GACATTTGAA ACGAAAAATTA ATAATGTTAG CTTAAATTTA | 600  |
|    | AAAGTACAAG ATGAGTATAA AGATTACTAT GATAAAAAAG TTGAAGCTGT TAAAAATTTA  | 660  |
| 30 | TTGGCTAAG CTAACACGGA TAGTAACAAA GACAATGTAT ATGTGAATTT CTTGAGTGTA   | 720  |
|    | CGCTCTGGAG GCAGCGCATT TAATAGTACT TATAACTATG CATCACATAT AAATCCTGAA  | 780  |
|    | ATTGCAAAAA CGATTAAAGC AAATGGGAAA GCTAGAACGG GTTGGCTGAT TGTGACTAT   | 840  |
| 35 | GCAGGATATA CGTGCCTGG ATATGATGAT ATCGTAAAGT AAATTATAGA TAGTAATAAA   | 900  |
|    | TAAGGATTCA ATAATGATAT TAAGACGAGT ATGAAAAATG TTAGATTCTA ATTATTTTCA  | 960  |
| 40 | CTACTCGITT TTATTTTGAA AATAAGTAAT AATTCAACAA TATTATAAAT TGAACAGATT  | 1020 |
|    | GTTTGTGAAA TTTTTGATA TATTAAAGTG AAAAAAGTGT ATAAATTGAT AAATATATGT   | 1080 |
|    | AATTAAACAAA AACAAATCAT TTTAAAGAAG AGAGAGTTGT AAGATGATGA AACGATTAAA | 1140 |
| 45 | CAAAATTAGTG TTAGGCATTA TTTTTCGTGT TTTAGTCATT AGTATCACTG CTGGTTGTGG | 1200 |
|    | CTAGGTAAA GAAGCGGAAG TTAAGAAAAG CTTTGAAAAA ACATTGAGTA GTACCCCTAT   | 1260 |
|    | TAAAAATCTA GAGGATTTAT ACGATAAGGA AGGCTATCGT GATGATCAGT TTGATAAAAA  | 1320 |
| 50 | TGATAAAGGT ACATGGATTA TAAATTCTGA AATGGTTATT CAACCTAATA ATGAAGATAT  | 1380 |
|    | GGTAGCTAAA GGCATGGTTC TATATATGAA TAGAAATACC AAAACAACAA ATGGTTACTA  | 1440 |
| 55 | mTAGTTCGAT GTGACTAAGG ACGAGGATGA AGGAAAAACG CACGACAATG AAAAAAGATA  | 1500 |

## (2) INFORMATION FOR SEQ ID NO: 362:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1216 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 362:

|            |            |            |             |             |             |      |
|------------|------------|------------|-------------|-------------|-------------|------|
| TACATGTTTC | GGATGCTACT | TTATTTAGTT | TGAAGGTGCG  | ATTATGGACG  | TTAGCGCAAG  | 60   |
| AAGTTTATCA | AGAATGGTAT | TTAGGATCGA | AGTTGTATGA  | AGATGTTGAA  | AAGAAAAATAG | 120  |
| CACGAACTAC | TTTTAAGACA | GGTTATATTT | ATCAAGAAAT  | TATTTTGAGA  | CCAGTAGATG  | 180  |
| AAGTTAAGGT | ACTTCTGAAT | GATTTAAAAG | GTGCTGGTTT  | CGAATTAGGT  | ATTGCAACAG  | 240  |
| GTGCTCCTTA | TACTGAGACT | GTTGTGCCAT | TTGAAAATTT  | AGGATTGTTA  | CCATATTTTG  | 300  |
| AAGCTGATTT | TATTGCAACA | GCAAGTGATG | TTTTAGAAGC  | AGAGAAATATG | TATCCGCAAG  | 360  |
| CACGACCATT | AGGAAAGCCG | AATCCTTTTA | GTTATATCGC  | AGCTTTATAT  | GGTAATAATC  | 420  |
| GCGATAAATA | TGAATCTTAT | ATCAATAAGC | AAGATAACAT  | TGTAATAATA  | GATGACGTAT  | 480  |
| TTATAGTAGG | CGATTGCTTA | GCTGACTTAT | TAAGTGTCTCA | AAAAATAGGT  | GCAACGTTTA  | 540  |
| TTGGAACATT | AACAGGTTTA | AAAGGTAAGG | ATGCTGCAGG  | TGAGTTAGAA  | GCGCATCATG  | 600  |
| CCGACTATGT | TATTAATCAT | TTAGGTGAAC | TTAGAGGTGT  | ACTAGATAAT  | TTGTAAATTTG | 660  |
| ATTGTTGTTT | GACAGCATAA | CTTGTAAGTA | ATGATTGAAC  | CAAAGGTTTC  | ATATTGAGTT  | 720  |
| ACAATGAAAT | TAATAATGAA | AAAATGCCAA | GAAGCAATGG  | AAGTAATCCA  | ATGCTTCTCT  | 780  |
| GGCATTTTGA | ATTTACATAA | ATTGTTTATG | ACTGTACCGT  | CAATTTCAGTT | GTGAAAATTT  | 840  |
| GATTGTATTC | ACCAACTTGT | TAAAGTTCAT | CAATTATATT  | GTTTGAAACA  | GGTGATCAA   | 900  |
| CGGATAAAAT | CATTAGCGCA | TCTCCGCCCG | CTTCAGTTCT  | ACCTAAAGTC  | ATAGATGCAA  | 960  |
| TGTTGATATT | GTATTTACCT | AACAATGCGC | CAGTTTTTCC  | TACCATACCT  | GGAGTATCAT  | 1020 |
| TATGATATGA | CACAATTTGA | TATTGATTTG | GCITTAAGTC  | TACAGAAAAA  | TTATTAAATTC | 1080 |
| TAACAATTCT | TGGAACGAAA | CCTGTAAAGA | CAGAAGCGCC  | AACITTTAACG | GAATCGCAAT  | 1140 |
| CGTTGTATAG | TTCTACCTCT | AAGTAGTTAC | TAAACCTGT   | CTCTGCTTTA  | TTATTTTCAA  | 1200 |
| TATTTAATGT | CACITG     |            |             |             |             | 1216 |

## (2) INFORMATION FOR SEQ ID NO: 363:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 867 base pairs

(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 363:

|            |            |            |             |            |            |     |
|------------|------------|------------|-------------|------------|------------|-----|
| ATCAAAATAT | TTTAAATAT  | ATTGATGGTC | ACATTGTAAA  | TmAATAGAA  | AATAAATTTG | 60  |
| ATCAAAATAT | ACAAGAACCA | TTAAATCCAT | TAAATTATGA  | TACTGTCACT | GGATTAGCTG | 120 |
| GGATAGGGAG | ATAATTGCTA | AATAGAGTAG | ATGAGAATGA  | ATTTAATGTT | AAAGCATTAA | 180 |
| AAAGCATATT | AGTATACTTT | AAAGATATTC | AATATTCTAA  | AAATAGCTGG | GTAGTCCAC  | 240 |
| AAGArAGTCA | ATTTTTAGAG | TCTGATAAAA | ATTATTTTAC  | TGAAGGTAAT | ATCAATCTTG | 300 |
| GCCTTGACA  | TGgAGTGCTA | GGACCGATGT | CTTTATTTGC  | ATTATGCCGT | ATTAAAGGAA | 360 |
| TTACGATTGA | AAATCATCAG | CACATATTAA | AAGACATGTA  | CAAATTTATC | ATGGACGAAA | 420 |
| AATTTTGTAA | CCACGAAAGA | TGGTTGCAGC | GTTACGATTT  | AATTTCTGAA | CGTAATCATT | 480 |
| TCAATTTTAT | TCGGAATGCT | TGGTGTATG  | GCAATACGGG  | TGTAATGACG | ACGTTGTTTT | 540 |
| TAATCGGCCA | AGCATTACAA | GATGATGAAA | TAATTAATAAT | GTCTAAAAAA | GTGATGCTAC | 600 |
| AAGTAGTAAA | TGATAAAGAT | GAAAATTTAA | TAAGTCCAAC  | TATTGTTCAT | GGATTGTCAT | 660 |
| CACAAATATT | AATGTTAACA | ATTATGAATT | TGAATTTTGA  | ATTAAATGAA | GTGTCTGATT | 720 |
| ATATCACTGT | ATTAATAAAT | AAACTGATTT | CTCATTATAA  | GGAAGATTAT | CTGGTGAATT | 780 |
| TTATAGACAT | TAATGAAAAT | AAGCAAGATG | TATTTAAAG   | TAGGAAAGTT | GGCCTTTTAT | 840 |
| AAGGTGAATT | AGGGGTCACT | TAACATT    |             |            |            | 867 |

35

(2) INFORMATION FOR SEQ ID NO: 364:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10813 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 364:

45

|            |            |             |            |            |            |     |
|------------|------------|-------------|------------|------------|------------|-----|
| TACCTTTTCT | TtTAAATCAT | TTTATATTTT  | CCCACTAATA | TCCGcTGcTA | ATCAATCCTG | 60  |
| ACATCctTGT | ATCActATGA | CAATTFAATTG | TTAAATACAT | GAATTTCTAC | ATTTTATGAA | 120 |
| AAAATCCATT | TTTATTACAA | TTCAACACTT  | TATATGACAA | CTTCATTACA | GTTACTTTTA | 180 |
| TTGTTGATTG | CTTACATTGT | TTTCTAAAAA  | AAATTGTGTA | TCATAATTAA | CGTTGAATAA | 240 |
| AGAAAAAAT  | TAACTTGGGA | GATAAAAAATG | GAATATAAAA | AGATACTAAT | TGTTTATTAA | 300 |

55

|    |             |             |             |             |            |             |      |
|----|-------------|-------------|-------------|-------------|------------|-------------|------|
|    | CAATCGCATG  | CCGACAGTTAA | TTATTATAGT  | AAAAACCAAT  | GTACATGGTG | GGCACTTTAAA | 420  |
|    | CGTCGCGCAC  | AAGTCGGTAA  | ACCTGTTTCT  | AATGATGGGG  | GCAATGCTAA | AAATTGGTAT  | 480  |
| 5  | TACAATGCAC  | GTAATCAAAA  | ATATGCGACT  | GGCTGTACAC  | CAAGAAAAAT | TGCTGTCAATG | 540  |
|    | CAATCAACTG  | CAGGATATTA  | TGGACATGTC  | GCAGTTGTTG  | AACAAGTATA | TAAAAACGGT  | 600  |
|    | AGTATTAAAG  | TTTCAGAAAT  | CAACTTTTAT  | CGCCCAATTAA | AATACAATAC | ACGTTGACTA  | 660  |
| 10 | AGCAAAAAGG  | CAGCAGCTAA  | CTTTAACTAT  | ATTTACTAAT  | CAAAAAATTT | CTATCACGAA  | 720  |
|    | CGCTTCAATT  | TCCTGTATGC  | CGTATGATGA  | AGTTTITTAIT | TATGGAAGTT | ATATTATTAC  | 780  |
|    | TTCTACAAAT  | TCAAATATGC  | GGTAATTGAA  | GCTATTATTC  | TTCTTCAACT | AITATTTCAT  | 840  |
| 15 | CTTTAGCATA  | ATCTATATAT  | AAAAATTTAT  | GCTATTATTI  | AAATAATTCG | CTATAACTTA  | 900  |
|    | ACATACGTTT  | TCGATATAAA  | CCTTGTTCTA  | AATCTCAATA  | ATTTTTTGCT | GTTTTCATCG  | 960  |
| 20 | TCATTAGTTA  | AAAAATAAAT  | TTAACTGAGT  | TTTTGAGACC  | TGTTTAAATT | GATGACTCAA  | 1020 |
|    | AACCTTTATC  | CCCTTTTCAC  | TCGGTTTAAT  | TGCTTTAATA  | TTTAAACACG | TCTCATTGTA  | 1080 |
|    | ATTTTGCTAC  | TAAGTTTGAA  | ATATTTCGAT  | TCGAATGTGA  | ACATCATTTA | TTATTACTTT  | 1140 |
| 25 | CGACAGCTAG  | AAAAATATTG  | TTAAACCAAA  | ACTATAATTA  | CCACTTTTAT | ATAGAACTAT  | 1200 |
|    | ATATAAATTT  | TAACTTTGAG  | ACAGAAGCTG  | TAGTCAGGTT  | TATGAATATA | TTTCTTTAGT  | 1260 |
|    | TTACTTGaTA  | TACTTATTGG  | TAAATCATTa  | TTTTATTGAA  | GAGCACTTAT | AACAAAAAGT  | 1320 |
| 30 | ATAATCCAG   | TTATAGCAGT  | TTCTGACCA   | TTTTTATCAG  | TAGAAAGTTC | ATCAACAAAA  | 1380 |
|    | ACTTTTGTAT  | AAATTTATAT  | ATTCATATAA  | TTTATTATGG  | TACGATATCT | ACCACGGTATC | 1440 |
|    | TATAACAGCT  | ACAGTTATCA  | TGCGCTCTTA  | TCCAAAATAG  | TTTTAAGAAA | TAAATATAAT  | 1500 |
| 35 | TCATTAAATAT | CATGGCTATT  | TGTAACGTCA  | GeGCTTaAG   | CCTTGTCAAT | TATTGTTTTA  | 1560 |
|    | TCTGTTTCTA  | AGAGCCCTG   | ATTTTCTGCA  | TACATATTTT  | GTATATCGCT | GTTAATCTCA  | 1620 |
| 40 | TTTAATTTTT  | TAAATTTTTT  | TCTAACCTTT  | TTCAAAATAA  | AATGCTCTTT | AATTCATCT   | 1680 |
|    | CTTAAATACA  | TATAAATCTC  | TATTATAATC  | TCCATAAATG  | TCATATCTCC | TCTCTCAAAA  | 1740 |
|    | AACATTTATC  | TCAATATATT  | AAAAATAAGA  | AAGGTTTTTT  | CAGAAATTTT | AGTATTTTCA  | 1800 |
| 45 | GATATTTTAT  | CTTAAGTATC  | TTATTATTAC  | TTCTTTTATA  | CGGTAGTGGa | ATAAATTAATA | 1860 |
|    | ATAATTGTAT  | TAAATTGCAA  | TGCTGTTGTT  | CTTTCATACT  | TACAAGCAAG | TCATAAGAAA  | 1920 |
|    | TGAGAAAAAT  | AATTTCTTATA | TGATGAATAA  | ATGATGACTG  | CAGAAGAAAT | TGCTAAAAAA  | 1980 |
| 50 | GGAGTGGAGT  | GAGGCATCTT  | ACTGTTTCATA | TTGTTAAACA  | ATATCAAAAT | TATGAGGAAT  | 2040 |
|    | TGAAGAAATCT | AAGACTGTTA  | AAATCGAGAC  | AAGAAGAAATC | TCGTTCCAAA | TTTGAATCTCT | 2100 |



|    |  |      |
|----|--|------|
|    | CTGTAAACCA AGACATCGTT TCAATGGGCG CTAACGCTGC TAAGGACATT ATAACAGACG  | 4020 |
|    | AAGACAAAAA GAAATTTGGT ATGGTAATTG TGGCAACTGA ATCAGCAGTT GATGCTGCTA  | 4080 |
| 5  | AAGCAGCGCG TGTTCAAATT CACAACCTAT TAGGTATTCA ACCTTTTGCA CGCTGCTTTG  | 4140 |
|    | AAATGAAAGA AGCTTGGTAT GCTGCAACAC CAGCAATTCA ATTAGCTAAA GATTATTTAG  | 4200 |
| 10 | CAACTAGACC GAATGAAAAA GTATTAGTTA TTGCTACAGA TACAGCACGT TATGGATTGA  | 4260 |
|    | ATTGAGGCGG CGAGCCAACA CAAGGTGCTG GCGCAGTTGC GATGGTTATT GCACATAATC  | 4320 |
|    | CAAGCATTTT GGCATTAAAT GAAGATGCTG TTGCTTACAC TGAAGACGTT TATGATTTCT  | 4380 |
| 15 | GGCGTCCAAC TGGACATAAA TATCCATTAG TTGATGGTGC ATTATCTAAA GATGCTTATA  | 4440 |
|    | TCCGCTCATT CCAACAAAGC TGAATGAAT ACGCAAAACG TCAAGGTAAG TCGTAGCTG    | 4500 |
|    | ACTTCGCATC TCTATGCTTC CATGTTCCAT TTACAAAAAT GGGTAAAAAG GCATTAGAGT  | 4560 |
| 20 | CAATCATTTA TAACGCTGAT GAAACAACCT AAGAGCGTTT ACGTTCAGGA TATGAAGATG  | 4620 |
|    | CTGTAGATTG TAACCGTTAT GTCGGTAATA TTTATCTGG ATCATTATAT TTAAGCCTAA   | 4680 |
|    | TATCATTACT TGAAATCGA GATTTACAAG CTGGTGAAC AATCGGTTTA TTCAGTTATG    | 4740 |
| 25 | GCTCAGGTTT AGTTGGTGAA TTTTATAGTG CGACATTAGT TGAAGGCTAC AAAGATCATT  | 4800 |
|    | TAGATCAAGC TGCACATAAA GCATTATTAA ATAACCGTAG TGAAGTATCT GTTGATGCAT  | 4860 |
|    | ATGAAACATT CTTCAACCGT TTTGATGACG TTGAATTTGA CGAAGAACAA GATGCTGTTT  | 4920 |
| 30 | ATGAAGATCG TCATATTTTC TACTTATCAA ATATTGAAAA TAACGTTTCT GAATATCACA  | 4980 |
|    | GACCAGAGTA GTCGGTGAT TTAACACACA TATAATAAAA CCTAAAAGCA CAGTAAGAC    | 5040 |
|    | CACCTTAAT TGAAATCGTC TTACTGCTGT TCTCTATTTA TAACACTTCG TATTGAATGA   | 5100 |
| 35 | ATTCATTATG CCTATTGAC ACATTATTGA AGTTTCTCTA ATGCTGGAT CCTTTATACG    | 5160 |
|    | TTACGGCTTC GTGCTATGTT TTGGTACATA AAGCTTTGAC ATATCGATAT TCTCCAACCT  | 5220 |
| 40 | TAACAGCTTA ATTTTATTAT TAATCGTTCC ACGAACCTCT GTTAAGCTAC CCGTTTACC   | 5280 |
|    | GACAACCGA TGACATGGCA CGATATAGA TAATGGATTA CTTCGACTG CACCTCCAAC     | 5340 |
|    | CGCTGGGCT GACATTTTGT GCTTGTTAAG CAGCTTGCT ACTTTTITGG CAATAGCACC    | 5400 |
| 45 | ATACGTTGTT AGAGTCCCAT AAGGAACCTG TCTTAATTCA TTCCAAACAC ACTGTTGAAA  | 5460 |
|    | ATGACTACCT GTTGGCTTTA AAGGTATTGT GATTTTCAGGA TTGTCACCTT TAAATACGC  | 5520 |
| 50 | GTCTAACCA TGTTGCGCCT CTCTAAATAT CGCTAAAGAC GTATTTTCTT CCCTAGTACC   | 5580 |
|    | ATCACCTTGT TGATTTTCAA ACAAAACAGC GGTGAGACTT ACCCATCATC TCAAAAGTTC  | 5640 |
|    | CAATCGTCCT ACAGGCGAAT CATAGTAACCT CTTATACTCC ATAAAAATTC CCCCTTTTTC | 5700 |

|    |  |      |
|----|--|------|
|    | ATAAGTCGTC AATTACGTAT ATAAACACGT AATACCAGCT ATCACTTTGC TGCAATATAC  | 5820 |
| 5  | AGTTACATAT CTTACTACAC GTGCTAACCT CTTACTTTGT AAACCAAATC TTAATTAATA  | 5880 |
|    | ATATTGAAAA TGCAATGAAT CCTTAATATT TTATTAAACC TATAATTACT TATTAAAAAT  | 5940 |
|    | AACACACAAT ATTCATAAAG TTTTAAAAAT ATTCTGTTTT ATCACCCTACT ATTAGTGGAA | 6000 |
| 10 | AAGTACAATT GCAATTGTAT ATAGTTTGCA TAACGCTTCA AAAGTAATTT CTTTTTTGTT  | 6060 |
|    | TAGTTCAAAA AAATTTAGAG GTGATGTTAT ATGAATAACG GTTTTTTCAT TAGCGACTTT  | 6120 |
|    | GATTCAATTT TTCAAGAAT GATGAAAGT ATGCAAGGTT CAAATCAAGT CGGAACACAA    | 6180 |
| 15 | AAGTACTATA TTAATGGTAA AGAAGTTTCA CCTGAAGAAC TAGCGCAACT CACACAACAA  | 6240 |
|    | GGTGGCAATC ACTCTGCTGA ACAAAGTGGC CAAGcTTTT CAAACAGCAG ACAAGACAA    | 6300 |
|    | CAAGGGCAAC AAGGTGGCAA CGGCAATTAT TTAGAACAAA TTGGTCGTAA CCTTACGCAA  | 6360 |
| 20 | GAAGCAGGTG ACGGTTTATT AGATCCAGTC ATTGGTCGTG ATAAAGAAAT TCAAGAAACT  | 6420 |
|    | GCTGAAGTTT TAAGTAGACG AACTAAAAAC AATCCTATAT TAGTGGAGA AGCTGGTGT    | 6480 |
|    | GGTAAAACTG CGATTGTTGA AGGTTTAGCA CAGGCAATCG TTGAAGGAAA TGTACCAGCA  | 6540 |
| 25 | GCAATCAAAG ACAAAGAAAT TATTTCTGTA GACATTTTCAT CATTAGAAGC TGGAAACGAA | 6600 |
|    | TATCGTGGTG CTTTGAAGA AAATATTCAA AAATTAATCG AAGGTGTTAA ATCTTCACAA   | 6660 |
|    | AATGCCGTAC TATTCCTTGA TGAAATCCAT CAAATATCG GTTCAGGTGC CACAGGAAGT   | 6720 |
| 30 | GATTCAGGTA GCAAAGGTTT ATCTGATATT TTGAAACCTG CATTAAGTCG TGGTGAGATT  | 6780 |
|    | TCTATTATTG GTGCAACAC ACAAGATGAA TATCGAAACA ATATTCTTAA AGATGCTGCA   | 6840 |
|    | TTAACGCGCA GATTTAATGA AGTGCTTGT TATGAACCAA GCGCTAAAGA TACTGTGAA    | 6900 |
| 35 | ATTTTAAAAAG GTATTGCGCA AAAATTCGAA GAACACCATC AAGTAAAATT ACCAGATGAC | 6960 |
|    | GTATTAAAAAG CATGTGTTGA CTTATCAATT CAATATATTC CACAACGATT ATTACCAGAT | 7020 |
| 40 | AAAGCAATCG ATGTGTTAGA TATTACAGCA GCACATTAT CTGCGCAAAG TCCAGCTGTC   | 7080 |
|    | GATAAAGTTG AAACGTAAAA ACGAATTTCT GAATTAGAAA ATGATAAAGC TAAAGCAGTA  | 7140 |
|    | AGTGCTGAAG AATATAAAAA AGCTGACGAC ATTCAAATG AAATCAAATC ATTACAAGAT   | 7200 |
| 45 | AAATTAGAAA ATAGTAATGG TGAACATACT GCTGTTGCTA CAGTTCATGA TATTTTCAGAT | 7260 |
|    | ACTATTCAAC GATTAACTGG TATTCAGATT TCTCAAATGG ATGATAACGA TATTGAACGT  | 7320 |
|    | TTAAAAATA TTTCTAATCG TTTAAGAAGT AAAATCATAG GTCAAGATCA AGCTGTAGAA   | 7380 |
| 50 | ATGGTTTCAC GTGCAATTCG CCGTAATCGT GCTGGGTTTG ATAGCGCAA CCGTCCAAAT   | 7440 |
|    | GGCAGTTTCC TATTTGTTGG CCTACTGGT GTTGGTAAAA CAGAGCTTGC TAAACAATTA   | 7500 |

|    |             |             |             |            |             |            |      |
|----|-------------|-------------|-------------|------------|-------------|------------|------|
|    | GACACACAG   | CTGTTTCAA   | AATGATTGGT  | ACAACTGCTG | GTTATGTTGG  | TTATGATGAC | 7620 |
|    | AATTCAAA    | CGTTAACTGA  | AAAAGTACGC  | CGTAATCCAT | ACTCAGTCAT  | TCTATTGAT  | 7680 |
| 5  | GAAATCGAAA  | AAGCAAAATCC | ACAAATTTTA  | ACATTGTTAT | TACAAGTAAT  | GGATGATGGT | 7740 |
|    | AATTTGACTG  | ATGGTCAAGG  | TAATGTCATC  | AACTTTAAAA | ATACAATTAT  | TATTTGTACA | 7800 |
| 10 | TCAATGCTG   | GCTTTGGCAA  | TGGCAATGAC  | GCTGAAGAAA | AAGATAATTAT | GCACGAAATG | 7860 |
|    | AAAAAATCTT  | TCCGCCCTGA  | ATTCCCTTAAC | CGCTTCAAGC | GCATCGTTGA  | ATTCTTACAT | 7920 |
|    | TTAGATAAAG  | ATGCATTGCA  | AGATATCGTC  | AACTTATTAT | TAGACGATGT  | ACAAGTTACA | 7980 |
| 15 | TTAGACAAA   | AAGGTATTAC  | GATGACGTT   | TCTCAAGATG | CGAAAGATTG  | GTTAATTGAA | 8040 |
|    | GAAGGCTATG  | ATGAAGAATT  | AGGTGCACGT  | CCATTAAGAC | GTATTGTTGA  | ACAGCAAGTA | 8100 |
|    | CGTGACAAA   | TTACAGATTA  | CTATTTAGAT  | CATACAGACG | TTAAACATGT  | GGATATAGAT | 8160 |
| 20 | GTTGAGGATA  | ACGAATTAGT  | CGTAAAAGGT  | AAATAACGAC | ACTTTAACAT  | ATCGCGCATC | 8220 |
|    | AAAAATGAGC  | ATCAGGTGCG  | CCTTGCGCTGT | GCTCAITTTT | TTAAITTTT   | CCCTGGAAAA | 8280 |
|    | TGATTGCTGT  | TGTGCTGTT   | TGTTCCACAA  | CAATCAGCAT | TAATGTCACA  | TGTACCACAT | 8340 |
| 25 | TTCTCTTGT   | TTGAACGCTT  | GAAAAATTTT  | ACTAGTGAT  | ATAAGGCATA  | TCCGAAAAAT | 8400 |
|    | GCTAAAAAAA  | TTAAAAATGTT | ATAATGACT   | GACACTTTAA | CCACTCCTTA  | AACAAATAAA | 8460 |
|    | TGTCGCACTT  | GATAAAAAAT  | GAATGTTAAG  | ACATATGCAG | TGACTAGAGG  | ATAGGCAACT | 8520 |
| 30 | GCAATGTCG   | TCCATTTCGA  | TGAATAAGTC  | TCTTTACGGA | TTGCTGCTAC  | TGTAGAAACA | 8580 |
|    | CAAGGAATAT  | ACAATAGTAT  | AAATATCATA  | AATGCATACG | CAGATAGCGG  | TGTGAATTGA | 8640 |
|    | TTTGAATCA   | CATTAAACAAG | GCCTGCATCA  | CCTGATGAAT | AGATAATCGC  | CATCGAACTT | 8700 |
| 35 | ACGATAACTT  | CTTTTGCTAA  | AAATCCTGGC  | ACTAACGTAG | CACCTGCTTG  | CCATGTTCCA | 8760 |
|    | AATCCGAGCG  | GTTGCACTAA  | CATACCAAAG  | AACTACCA   | CCATATGTAA  | AAAACCTTGA | 8820 |
| 40 | TTGATATTCA  | CATTGATACC  | ATGTGGTCTT  | ACATAACTTA | ATAGCCAAAT  | GACTACTGAG | 8880 |
|    | CGCCAAAAA   | TAAATGTACC  | TGCTTTACGA  | ACAAAGCCCT | TAGCCTTTTC  | CCAAGTACTA | 8940 |
|    | CGCCACAACG  | TTTAAATGGA  | AGGCACACGG  | TATGTTGGCA | ATTCACAAT   | AAAGATTGCA | 9000 |
| 45 | TTATCATTTT  | TTAAAAATCGT | CTTAGTAAGT  | ACTGTACTGA | CTAAAAATGC  | CATAATAATA | 9060 |
|    | CCTAAAAACAT | ACAGGCTTAA  | TACTACTAAA  | GATTGATTCT | CTTTGAAAAA  | GATACCTACG | 9120 |
|    | AACAACGCAT  | ACACTGGCAG  | TCTAGCAGAG  | CATGACATGA | ATGGTGCAAT  | TAATATCGTT | 9180 |
| 50 | GTTAAAGCCT  | CTTTTTCATT  | TTCAATACTG  | CGCGCAGCCA | TAATACTCGG  | TACATTACAA | 9240 |
|    | CCAATCCGA   | TAATCATTGG  | TATAAAAGAC  | TTCCCGCTTA | AACCGAACGA  | TTCCATAATA | 9300 |



AAAAGAGCA CAACAATITG TGGTACAAAG ACTAATACTG ATCCTACACC AGCAATAATG 9420  
 CCATCTGTAA TTAAATCTTG TAAAAATGGT ATAACACCAA GATAATTCAT AATCGTCTTC 9480  
 5 ACCTATCTG TAAATGTACC ACCTATAAAT GCATCGAGTT GATCCGACAA AGGTGTGCCA 9540  
 ATCCATGTAA ATGTAGTTTG AAAGATCAAC CACATAATTG CTAGAAAGAT AGGCATCCCT 9600  
 ATATATTTAT GTGTTAATAT CTTGTCTATT CTAGAGCTGA AATAITGCTT ATCTTCATCT 9660  
 10 GGATACGTTA CCACGCTCTG CAATAACGTC TCAATATAAT GATTGCGTAT ACGTCCATC 9720  
 TCTCGACGAA CAGATACAGC CCTTACTTGT TCAGCAACTT GATCAGTAA ACTCGACAAT 9780  
 TTATTTACAA CCTCTGAATT AAGTTCGTTT GCAATTGCA TGTATTTTAA TAAGAATTGA 9840  
 15 ATGCCAATAA ACCTAGCTTG ATACTTATCA TGAGATGTCT CTGTCAATTAT TATTTGACAC 9900  
 ATATTTTATA TTGTCTCTTC AATCTTCTCA CCATAATTGA TTTTAAATG CGGTGATAC 9960  
 20 CCTTCCCTTA GATGCTTIAT TTGCGCAAGT AAATATTTTG TTCCTTTGCC TGTACGTGCC 10020  
 ACAACTGGAA AAATAGGTGT TTTTAACTTT TTCATCAATT TATGATAATC GATTTTATC 10080  
 CCGCGCTTG TAGCTACATC AATCATATTT AATCCGATGT ATATTGGTTG ATTAAGTTCT 10140  
 25 AACAAATGTA CTGTAAATG CATATTTCTT TTTAGTTGAC TCGCATCAAC AATGTTAATG 10200  
 ATTCTGAAA ATGAATCGTT TAATAAATAG TCTGTCACTA CAGTTTCTT TTTAGAAATC 10260  
 GCGATAAAT CATATGTACC TGGTAAATCA ATTAATTGTC CTACATTTTC TTTAAGTTTC 10320  
 30 CCTACTTTTT TCCTACCGT TACGCCACTC CAGTTGCCTA TATATTCTA CGAACCAGTT 10380  
 AAAGGGTTAA ACAAGATGT TTTACCAACA TTAGGATTTT CTAAATACA ATAATTTTCC 10440  
 ATTCTGTCGG CTCCTATTCT TCTAATGCAA TAGAACAGC ATCGCAATGT CTAATACTTA 10500  
 35 ACTGTGTCC GTTTACTTCA ATAATACATG GCCCTTTAAA TAAACATTTT TGTTTAATCG 10560  
 TTATGATAGC GTCATCTGTT AACCCAAAGG CACTTAGACG ATACAACATA TTCTCAITAG 10620  
 40 CAATATCCAT TCGCTTIAIT TTATAAGCCT TATTCATTTT ACCATTTTTA ATGTTTAAAC 10680  
 TACTATTTCG CTCTCTTATT AGAAATAATA ATCATTATCA CTTAAAAATC ATAACCTTTA 10740  
 AAATGTAGC TCGCAATACT TTATTTAAAT AATTTTCATT TTTCAITGTA AATTTGTGAC 10800  
 45 ATTGCAAAA TGT 10813

(2) INFORMATION FOR SEQ ID NO: 365:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 6804 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

EP 0 786 519 A2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 365:

|    |  |      |
|----|--|------|
|    | TSATGGATTA CGACACATTT TACGAGCGAA TGGTTTCAAA GTGTTTGGTC CAAAATAGCA  | 60   |
| 5  | AGCAGCTCAA ATCGAAGGCT CAAAATTATT TGCTAAAAAG ATAATGGAAA AATATAATAT  | 120  |
|    | TCCAACGTCT GATTATAAAG AAGTTGAGCG AAAAAAGGAT GCTTTAACAT ATATTGAAAA  | 180  |
|    | CTGTGAATTG CCGTTTGTG TCAAGAAAAG TGGGTTAGCT GCTGGGAAAG GCGTTATTAT   | 240  |
| 10 | TGCAGATACT ATTGAAGCAG CCAGAAGTGC TATTGAGATT ATGTATGGTG ATGAAGAAGA  | 300  |
|    | AGGTACTGTT GTATTTGAAA CGTTTTTAGA AGGTGAAGAG TTCTCGCTAA TGACATTGT   | 360  |
|    | TAATGGTGAT TTAGCAGTAC CTTTCGACTG TATTGCACAA GATCATAAAC GCGCATTTGA  | 420  |
| 15 | TCATGATGAA GGACCAAATA CTGGTGATAT GGGGGCTTAT TGTCGmGTAC CACATATTAG  | 480  |
|    | TGACGATGTT TTAATACTTA CAAATGAAAC AATTGCACAW CCCATTGCAA AGGCAATGCT  | 540  |
| 20 | TAATGAAGGT TATCAATTCT TCGGTGATT ATACATTGGT GCTATTTTAA CTAAGATAGG   | 600  |
|    | TCCAAAAGTA ATAGAATTTA ATGCCCGTTT TGGTGATCCT GAAGCTCAAG TATTATTAAG  | 660  |
|    | TGCGATGGAA ACTGATTAA TGCAGCATAT TATTGATTTA GATGAAGGAA AACGTACTGA   | 720  |
| 25 | ATTCAAATGG AAAAATGAAT CTATTGTAGG GGTCAATGTT GCATCAAAAG GATATCCTGA  | 780  |
|    | TGCATATGAA AAAGGCATA AAGTAAGTGG CTTTGATTTA AATGAAAATC ATTTTGTTAG   | 840  |
|    | TGGATTAAAG AAGCAAGGTG ATACCTTTGT TACTTCAGGT GGTAGAGTTA TACTTGCCAT  | 900  |
| 30 | CGGAAAAGGT GACAATGTAC AAGATGCACA GCGAGACGCA TACAAAAAAG TATCACAAAT  | 960  |
|    | ACAAAGTGAC CATTTATTCT ATCGTCATGA CATTGCGAAT AAAGCACTAC AACTTAAATA  | 1020 |
|    | AGTAAATTTA AAATACTAAG aTTAGCTATG AACGAATCTA TAACGATAGA TTTTTCATA   | 1080 |
| 35 | GCTTTTTTAG TTGTAGAGTC TAGGACATTG ATTTCTGTAC CAAATTGTGT ATTATGCATA  | 1140 |
|    | TGTAATACAA AAGAGGCGCC ACAACATGTT TGGATGAACA AAATAACATG TTTGTGGCAC  | 1200 |
| 40 | CTCTTTTGTT TAGTATGGAA TAAATGGTTT TCTTTTCTA TACAATGAAT TTCTAATTTA   | 1260 |
|    | GTATCTATAC AATTATGGAT AAAATTTTAA CTAACGAGCC AAGACGAACA TCATCTATGC  | 1320 |
|    | CCGTGATGGG TAAGGTGATT GAACAATAAT ATGCCATAGT AATAATGGCA ATTAAGAACTA | 1380 |
| 45 | TAATAAGAT TATATCTTTA TATGAGAAAG GTACGTTGTA ATAGTAAGTA CGAGGACCAT   | 1440 |
|    | CTCTAAATCC TTTGCACTCC ATCGCAACTG ATAATTGATG TGCCCTTCTA ATATTTTGGC  | 1500 |
|    | TTAATAGAGG TATAATTAAA TGCTTAAATC GCTTTAACCC TCTATAATTG GCCGCGTCTA  | 1560 |
| 50 | TCATCTGATA GCGCATTTTT AAAGATCTGC GAAGcTGTA TAAAGAACTA ATCATTAAAG   | 1620 |
|    | GTATCATACG AATGGCAGCC ATGAATGCAT AAGCAACTTT TGATTTTAAC TTTAAATGTT  | 1680 |

55

|    |  |      |
|----|--|------|
|    | ATGAAATAGC AATGGTTCTT AATGATACAT GTAAACCACG AACTAAACTT TCTGTTGTA   | 1800 |
|    | TATGGATAAA TCCGAATTC AAAATTGTAT GGCTACCATT CCGGTATAAA ATCATGAACA   | 1860 |
| 5  | GGGAAGAGAG TAATGCAAG CCAATACTTA TAGTTATAAA AATTGCTGTA ATTTTAAACT   | 1920 |
|    | GAGTACCATT AAACATCAAT AAGAAACTA ACATTAAGAT AGTGATATA AGCATAAAAT    | 1980 |
|    | CGAAATTATG CACAATATA ATAAAGAAAA ATAGTATAAT TCCAAGAAAT AGTTTCGTTA   | 2040 |
| 10 | TAATGTTGAC ATCATCAACA AATGATTGCC GAACTTTCCA TTGCTCATAC ATTCGTATCA  | 2100 |
|    | CCATCACAAT CTAGTAACGC ACCATCTGAA ATTTTAAGTC TTCTTGATGG ATAACGTTCA  | 2160 |
| 15 | ATTATTTTCAT CGTCATGTGT AaCCATGACA ATACTTTTGC CCAATTAAT TCGCTTTTGG  | 2220 |
|    | AAAAGTTTGA TCAACTGGAA TGTATTATGG CTATCAAGTC CAAATGTCGG TTCACTTAAA  | 2280 |
|    | AAGATAATAT CAGCTTTAGA ACTTAGTGCG GTAgCTACGC TAAGGCGTCG TTTTGTACCA  | 2340 |
| 20 | ATAGACAACAT CATAAGGATG TTGATCTTTT ACATTTTGTG AATCTAAAAG TTTTAAAGT  | 2400 |
|    | TGTATCGTTT CATCATCACT TTGATCTTTA GAAAGGTGAT TAAATGAAT GTTAATTTCA   | 2460 |
|    | TCATAAACCG AATTGTTTAT AAATTGTAAAT TCTGGGTTTT GATAAACTAG GTACATGTGT | 2520 |
| 25 | TTTGCTGCAT GTTTAATTTT TGTTAAACGC TGATTTTCAA AATAAACATC ACCTTGATAT  | 2580 |
|    | TTAATCAATT GCATAATTGA TTCAAGCAAG GTTGTTTTAC CACTACCAAT TGCCCTGTGA  | 2640 |
|    | ATTGTAATCC ACTCACTAG ACCAATTTCT AAATCTGAGA ATGAGAGCAA TGTTGATTTA   | 2700 |
| 30 | CCGCGAATAA TACGTCCATT TTTAAATTGT AATAAGTGTG AGTTTGTGTG TGGAAAGTCA  | 2760 |
|    | ACACGACITG GTGCGAATTC CCATGCACGT GGATGCCACA CACCATATTC ATCGAGTAAA  | 2820 |
|    | TGAACATACT TCTGTAATAT GATTTTCAGGA CATTCATCGG CAATGATATT TCCGTTATAA | 2880 |
| 35 | TCCATCAAAA TGAACGGTC GACATGATTC CAGATGTGTT TAACTTTATG TTCAACGATT   | 2940 |
|    | ACAACCGTTT GATCTTCCA AAGTTCAATT AGTTTAGTCC ATAAATCTTC TGTGCTTGA    | 3000 |
| 40 | ACATCTAACA TTGCTGTCGG TTCATCTAAA AACAATGTTT TIGATTGTTG AAGAAAGTGT  | 3060 |
|    | TCAACAATTG CCAATTTCTG TTTCATCCCG CCACTTAAAT CTTTGATATA CGTTTCAGGG  | 3120 |
|    | GTAACATTTA AATTGACCAT ATTTAAAGCA TTGATAATTA ACGCATCCAT GTCTTCACGT  | 3180 |
| 45 | GGTAATTGTC TATTTTCTAA AAGCAATGCA AGTTCTTCGT ATACTTTTGG CATACAAAAC  | 3240 |
|    | TGGCTATCAG GGTCTTGGAA AATAACGCCA CTTAATGGGT CAACGATTAG TTCAATCAT   | 3300 |
|    | TTCATAGGTA ATTCAATTAA ATTAGGAACA ATACCACCTA ATACATTCAG AAGTGTAATT  | 3360 |
| 50 | TTACCGCAAC CAGAAGACC GAGTAAAGT ACTTTTCTTT TGTCTTGAAT AGTGATATTT    | 3420 |
|    | AAATGATCGA AAATTTTACG TTGACCACTT GGATATTTTA ATCGTAAATC ACTTACTTTT  | 3480 |

55

## 55

1257

TTTTCATTAT TCCATTTACA ATTCATATAG GATATATCGT CTATAAAGAT GAATTGAATC 5400  
 AGGAAAATGT AAAAAATCAT TTCATGTGGA TAATTGTGAG TTATGGTATA AGTTACTTTAA 5460  
 5 TTACACAAAT TGCATTGTAT GGCAGAAATG ATGCTAATGA AATAGAGTCA ATTGATATCT 5520  
 TAAAGTGTCAA TGCTTTCTTT ATAATTATGT GGTACTTGG TCAAAATGGCT ATTTGGAAAT 5580  
 TCITGTTCCT GCGCCGAGCT TTACCTTTAA CAAAGCAAGA ATTAGGTGAA GAGGAGCCAG 5640  
 10 AATTATCAAG AACAAGTAAA GGGAAATGCA CGAATCAAAC TAAAAATCAC TTGAAACAAC 5700  
 TCCAAGATAA GACTACAGAA TATGCACGTA AGACAAGAAG AAGTGTGAT TTAGATAAAA 5760  
 TTAGAGCTAA AAGAGATAAA TTCAAAAAGA AAGTTAATGA TATTATCGAT ATTCAAGAAG 5820  
 15 ACGATATTC TGATTGGATG AGAAAAACCA AATGGGTTAA ACCAATGTAT GTCGAACTAT 5880  
 TTTGTGGTGT CGTCATCTTT TTATTCACAT TTTTAGAATT TAATAATCGT AATGCATTAT 5940  
 20 TTGTATCTGG TGATTGGAAA TTATCACAGA CACAATATGT TATTGAATGG GTTACATTAT 6000  
 TAATTCTGTT ATTCATTATT ATCGCATATA TCGCTACAAC GTTAACCTTC CACTTGAAAG 6060  
 GTAAGTTTTA TTATTTACAA TTATTTATGG GGAGCATTTT ATTCTTTAAA TTGTTAACGG 6120  
 25 AATTTATAAA TATAATGATT CATGGACTAT TACTTTCAGT GTTCATTACG CCAACATTAC 6180  
 TAITAATGTT ATTGGCAATC ATCATTTCTT ATTCGTTACA ATTACGAGAG CGACCAATAT 6240  
 TAAAAGCATT ATAAAAGTAC TATCTATTAA ACATTITGAT GTGTACGCTA TAAGTTAGAT 6300  
 30 ATATCTCTAA CTTACTTAGA TACAGGTCAA TGAAGTTTAT GGATAGTACT TTTTTTGTA 6360  
 CTAGATTGTA TTGATTCAGG TGATGTGAAT TAAGTATTGA TAATTGTATA CAAAGTTTTA 6420  
 GTGCAAAATA AATAGTTGAA AAGTTATCCA TTTGTAAAA CAAGAAAAC AGTAAATAGT 6480  
 35 TGAAGCGACT TATGGAATTT GCGAAACGAT ATATAGTATT TCCTTTGTAG AAATTTTACA 6540  
 TATATCATTC AAATTACTAA TTTGTTAAAA TCAACAGTAA GATTAGAAGT AGATGATATT 6600  
 40 GAAATTTGGC AAACAATTATA ATCTATATAA AACTACAACG AAACACAGAA AGGAAAGTTGT 6660  
 CAGATGAAAA TAGCAACTCT GAACAAAGGC AAAGAAACAA AATATTTTAA TGGATATCCT 6720  
 TTAATTGAAG AAGAGGATAT CTATTCACAA GATCATTTAA AAGAAGGAGA TATTTTTCAA 6780  
 45 ATTGCTGACTG ATAATCACA ATAT 6804

(2) INFORMATION FOR SEQ ID NO: 366:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1717 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 366;

55

## (2) INFORMATION FOR SEQ ID NO: 367:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1847 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 367:

|             |            |             |            |             |             |      |
|-------------|------------|-------------|------------|-------------|-------------|------|
| AAGATTAGGC  | ATCAAGACAG | GGTCGCGATT  | GTTTGAAATC | CCACATAGnA  | ATGATATTTA  | 60   |
| CaTTATCAAT  | CCaAGTATGC | GTAATATCT   | TAATGTTTCA | GTTGCTATT   | CTAAGATTGC  | 120  |
| ATTGCGTAT   | ATTCCACCTG | AAGATTTACA  | CCAATATAGT | ATTGACGAAT  | TTTTTATGGA  | 180  |
| TGTTACTGAT  | AGCTATCATA | GATTIAGTTC  | TACAGTACAT | GCATTTTGGC  | AAAGACTTAA  | 240  |
| ACGTGAAATT  | TATGAAGAAA | CAGGCATTTA  | TGTACTGTG  | GGCATTGGTT  | CTAATATGTT  | 300  |
| ATTAAAGTAAA | ATTGCTATGG | ATGTTGAAGC  | GAacATAGTc | AAAAATGGTAT | AGCTGAATGG  | 360  |
| CGATATCaAG  | ATGTACCAAC | GAAATTATGG  | CCAATTcMGC | CcLTGCAGAGA | TTTTTGGGGT  | 420  |
| ATTAACTGCTC | GAACAGAAGC | CAAAATTGAAT | AAAAGAGGAA | TTTTTACTAT  | AGGAGATTTA  | 480  |
| GCGAATATC   | CATATAAAAT | TTTAAAAAAA  | GAGTTCGSTA | TTTTAGGTGT  | TGATATGCAT  | 540  |
| CTACATGCGA  | ATGGGATAGA | TCAGAGTAAA  | GTACGTGAAA | AGCACAAGAT  | CAGCAATCCA  | 600  |
| TCGATATGCA  | AAAGTCAAAT | ATTAAATGAGA | GATTATCATT | TGTATGAAGC  | AAAAGTAGTA  | 660  |
| ATGCAAGAGT  | TAATTGAAGA | TGTTGCTAGC  | AGAGTTCGAG | CAAGAAAAAA  | AGTGCCAAGA  | 720  |
| ACGATACATT  | TTGCCCTTGG | CTATAGTGAT  | GAAGGCGGTG | TACATAAGCA  | ATATACCTTG  | 780  |
| AAAGATCCAA  | CAAACTTAGA | AAAAGATATT  | TATAAAGTAG | TAATGCATTT  | CGCAGATAAA  | 840  |
| TTATGTAATA  | AACAAGCACT | ATATCGTACG  | CTAAGTATAT | CTTTGAGTCA  | ATTTATTAAT  | 900  |
| GAGGATGAGC  | GACgTTAAGT | CTGTTTGAAG  | ATGAATACCA | ACGCAAAACGT | GACGAATGTC  | 960  |
| TAGCTAAAAC  | GATAGACCAA | TTACATTGTA  | AATACGGCAA | AGGTATTGTG  | TCCAAAGCAG  | 1020 |
| TATCGTTTAC  | AGAAGCAGGT | ACAAAACACG  | GCAGATTAGG | TTTAATGGCT  | GGACATAAAA  | 1080 |
| TGTAATGACT  | ATACGGTTTA | AGTAATATAT  | AACTGTGATT | CGTATAAAAT  | AAGTCTCTAA  | 1140 |
| AGATAAATAT  | TTCAATATAT | ACAATAGATT  | TTCACAATAA | TATCTAAGAA  | TACATGGAAT  | 1200 |
| TTATCAAAAG  | AGACTTAATA | ATTATTGGAT  | ATAACAATCA | AAATCACTCA  | ATGCTTGCAAT | 1260 |
| ACCGCGTTCT  | CGGTCAGTAG | GGTTTTTGAA  | ACTAATTTTT | AAAGCACCGCT | ATATATCTTC  | 1320 |
| CGCTACTTCT  | AAGATTCTTA | AGTTGCTTAT  | AGATATGTTA | TGTAAACTCA  | GGATATAAGT  | 1380 |

|    |            |            |            |            |            |            |      |
|----|------------|------------|------------|------------|------------|------------|------|
|    | TAGTCCACCT | AGTTGTTTAG | CGGFTAGTGC | GTCGCGATAC | GATTTAGCTT | GGGCAAAAAA | 1500 |
| 5  | TGATAACAAT | TTTTCAGAAT | CATTGCTTTC | AATTAGTCTC | TCCTAAATCT | GAAACTGACT | 1560 |
|    | TTTTAGCTGT | CGAATCATT  | CTAAAATATA | CGTTTATTA  | CTCAAGGTGA | TATCTTTCCA | 1620 |
|    | CATTGTGCA  | TTACTACTAG | CTATACGAGT | GATATCAGCA | AAACCACCAG | CTGCAAGTTT | 1680 |
| 10 | ATTAACATA  | TGATGTTCTT | GACCGTCTTT | TTGACTAACA | TGAACATAAC | TAGATGCAAC | 1740 |
|    | GATATGAGT  | AAATGACTTA | CGACGCTTGT | TGCGTAGTGC | GTTTCTTCAG | CAGTAGTATC | 1800 |
|    | AATAAATTTA | GCAAGAGTAG | GTGATACACG | TCTTTTAAAC | GTGTTTTCG  |            | 1847 |

(2) INFORMATION FOR SEQ ID NO: 368:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 494 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 368:

|    |            |            |            |            |            |            |     |
|----|------------|------------|------------|------------|------------|------------|-----|
| 25 | AACAAAGGC  | ATAAGTTCGT | GAATTAATGC | GTATACAAGG | ATAAAGCTTA | TAACAGTAGT | 60  |
|    | AATGTGTTGT | ATCAACGAA  | CACATATAT  | TCTATTTTCA | GATAGCAATC | GATTCACTAA | 120 |
| 30 | TCGATAATT  | ACGTATACAA | GAATTAACAA | CAGCAACAAT | TAAACAATA  | TCATATTGG  | 180 |
|    | CCCCATTAA  | TTTTTAAAT  | TGTTTTACA  | TCATTTCTA  | CTTTATTATC | ATACTAATTT | 240 |
|    | TAAAGGCAAA | GGTGGACATC | GGCAACCTCT | CGTAAACTAT | TATCAAAAA  | TAAACGTATC | 300 |
| 35 | TCATTGTTAT | GATATTTATA | AATCAATTCG | TTTTTATTAA | AGTCTTTTTT | AACAAGTTTG | 360 |
|    | TCACACTCTA | TTAAATAATC | ACGCAATGAT | CCTTTTAAGA | AATCATCTTT | ATAACTTGGT | 420 |
|    | GTGTACCAAT | TTCCATCTTC | TTCAATGACA | ATGTTGCCAA | TATCAAAATC | AAGGACCTTG | 480 |
| 40 | CCGTCTCTG  | AAGT       |            |            |            |            | 494 |

(2) INFORMATION FOR SEO ID NO: 369:

(i) SEQUENCE CHARACTERISTICS:

- SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 2518 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 369:

GTCCATATTT CCGGTCCAAC CAACTAAAAA TGGGTAAATC CCTGGATTTa AATCTACTTC 60



|    |             |             |            |            |            |            |      |
|----|-------------|-------------|------------|------------|------------|------------|------|
|    | GATAATTCTA  | AATGTAGCGA  | CCATCATTAC | GAGATGATT  | GATATAAGCA | CAGTTTGGAT | 180  |
|    | GTTGACCAAT  | ACTATCGCCT  | TCTTCTTCGA | TGATATCTAT | TTTAATACCA | TCATCAGCTG | 240  |
| 5  | CAATTTCTAA  | TGAAGATTTA  | ATTGCGTTAT | CAATGTTGA  | ATATCCCATT | GCTCCACCCA | 300  |
|    | CAATAGCGAC  | ATCTGTACCA  | TGTCCTTGGT | GTGTTTGAGC | AAATGATTCA | TAATAATGTA | 360  |
|    | TTTCAATATT  | TTTAGGAATA  | TCTCCCAATA | TGCGCGTGTC | TGAATTCCTA | ATCTTTACTG | 420  |
| 10 | CACCAGCCGT  | ATGAGAACTT  | GAAGGGGCCA | TCATAACTGG | TCGATAATA  | TCGAAAGCAC | 480  |
|    | TTTGATAATC  | ATAGCTCTTT  | GCCATAATTA | AACACTCTCC | TTAATATGAT | TCTTTTGTCT | 540  |
|    | GGGCATTTTA  | AAGTTGATAT  | TCATTAAATT | AACTTTATTA | ATTAGTGTTC | CAATAATATA | 600  |
| 15 | GGCTAAGACA  | ATCTGACAA   | TAATCACTGT | TGCGATTGTT | ACAATTGATG | TCACTGCATT | 660  |
|    | ATTAAAGCCA  | AACAATACGA  | TGCGCGCTGC | AATTGCTGTT | GCCATACCTT | TGACACCTAT | 720  |
| 20 | TACTAGTCCG  | CTAATGTCA   | CGATACATGC | GTGACGACC  | CCAATCAGTG | CATTGTGACC | 780  |
|    | ATATAGTTGT  | ACTGGATATT  | GCGCTATTAA | ATCAATTGTC | GTCAATGGCT | CAATACAAAC | 840  |
|    | TGCAAAATGCT | TTTGACGGTC  | CACCAATGTT | TAATTTTCGG | AATAAAAAAT | GGTTAAACAA | 900  |
| 25 | TGAGCTACCT  | GTACATGTTA  | GTGCTCCAAT | AGCCATAGGA | ACACCTGTCA | GTCTTAATAA | 960  |
|    | ACTTGTTAAT  | ACCAATTGAAC | TTAGCGGTGT | CATACCTGTA | ACAGGAATCA | CTAGCTCTAA | 1020 |
|    | AATGACCGCT  | AATGCATATG  | GATTGTTATC | ACCTACCGCA | GTGACAGCAC | TACCTATTGT | 1080 |
| 30 | TTTAAATGTT  | GCTAGCACAC  | CAGGTGTAAT | GATTGATGCA | AGTCCGAAAG | CAATTGCTGG | 1140 |
|    | TGCAAAATAG  | ATCACCACAA  | TTAAGTCCAA | GCCTTCTGGA | ACTTCTCTTT | CAATCCATTT | 1200 |
|    | AATTAAAAAA  | GCTACGCCAT  | AAGCTGCGAT | GAATGCTGGT | AATAATTTAA | AGTCATGTAA | 1260 |
| 35 | TACTAAACCA  | ACAATGACCG  | CAATACTGG  | TGCAACGCTC | AAGTTTAAGC | ACGTTAGAAT | 1320 |
|    | ACCTACTGCG  | ATACCGCTTA  | AACCTCTCTG | TAAATCCCCA | ATATCTTGTA | GAAATTTAAT | 1380 |
| 40 | ATCAAAATAG  | CCACCAATAG  | CATAACTTAA | GAATGCTTGT | GGTAGAAATG | TCGCACAAGC | 1440 |
|    | TGCACCGGAT  | AATGCTTGTA  | GTCTTGTGTT | ACCGTACGGT | GCATCTTTA  | AAAATAGCGT | 1500 |
|    | CATGATCACT  | AAAACCAAGA  | CTAATGTGCC | TACACCTAAC | AGAATATCCA | TTTCCCAAAA | 1560 |
| 45 | ACCTCTCTCT  | ATGTTTATTT  | TATTTTCAGA | CCATAAACAT | CGTACACCCT | TAAGAAAAAG | 1620 |
|    | TTTTCAACTT  | TTATCTGTTA  | TCAAATCAAA | TATTTAAGTG | AAATATTTC  | TATTTGTGAA | 1680 |
|    | AGATTTTATA  | AATGGATTGT  | TTCAAAAAC  | ACTTAGTTTG | TCGTAAATAT | TTACTAATTA | 1740 |
| 50 | ACTTTTACT   | CTATATTICA  | AACAGTTGIG | TGACAGTTT  | TTGATAACTT | TTTATACCT  | 1800 |
|    | GAAAGTAAGT  | AATATTICTA  | AAAACCTTTA | ATATTATATC | ACTTTATCTT | TCGAGCTATT | 1860 |

ACACATATAT TTGCAATAAG ATAATTAAAG TAGGATATTA TTTTAGITTT TCTGATAGGA 1980  
 ATGATGATAG TCTATAGGTT GAATCTTTAC TTTTITTTAA AGCTAAATTT ACATCAACTT 2040  
 5 AACAAATGGTT GGTATACTG AAGATGAAAA TATTAGTAT AACTTAGTGG AGGCGATAAA 2100  
 GGTGCAATTG AGTCATTCCG TTAAAGTTGC AATTCTATC TATTAGCAC TTATCTTTAT 2160  
 AAGTTCACT TCTATTITAG TCATTATTTT ATATACGAGT ATGACTGGAC ATGATGTATC 2220  
 10 ACATTTCGTG TTAGATAGTC AGCATTCTCA TCATGGATCT CTTACGCAAA AACATTGAG 2280  
 TCTTCTGAA ATCTCATTTA AATAGTTTAT CCTCTGTGTT TCAACATTCA TTTCCCATAT 2340  
 CGATTCATT ATCTATCATC TAGACCACTA CATCTAGAT GATTTTTTTA TTTTCTCATT 2400  
 15 TCACCTCTTC TTAAAAGTCG ATATAATGAA TTAATCATT ATCATAACC GACATATTTT 2460  
 ATGTTGTTGG TGTAAAGTTT AAAGGGGTGA GATACTTGGC GAATaATCaT TCAGCTTT 2518

20 (2) INFORMATION FOR SEQ ID NO: 370:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 790 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 25 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 370:  
 30 ATACTACTGG ACGTCnACC AATTGACAAA ACCTGAAAAC TGGATTTTTA TTGCAGGATT 60  
 ACATCGAACC TGAAGAACA ATCCCCAGGA GTTGATGATA TTCGAGATAT GGCATACAAT 120  
 CAAGGTAGTT TAGATAAGAC AATTTATGAA ATTTCTAAAC GCACAGTACT ATTTTAAATA 180  
 35 CAGAAAGATA TTACGGTATA TAATAAGACG ATTGACTGTT TAAAITAITA TAACTATAGT 240  
 GACGAAAGAA TAAAGGATGA TTAATGAAT TCACAAGAAT TATTAGCAAT TGCTGTGGAT 300  
 40 GCAATTGACA ATACCCCGAG CGAAGATACG ATTTCTTTAG AAATGAAAGG TATCAGCGAT 360  
 ATGACAGATT ATTTGTTGT AACGCACGGA AATAATGAAC GACAAGTTCA AGCGATTGCT 420  
 AGAGCGGTGA AAGAAGTAGC CAATGAACAA AATATAGAAG TAAACCGTAT GGAAGGATAC 480  
 45 AATGAAGCGC GTTGGATATT AATTGACTTA GCTGATGTTG TGGTACATGT TTTCCATAAA 540  
 GACGAAAGAA ATTATTATAA TATTGAAAAG TTATATCAAG ATGCACCATT AGAATCATAT 600  
 AGTCAGGTG CTTATTAAAT ATGTGCAAT ATGCAGAAAT GAGCCTAGTG TACGATCAAT 660  
 50 TGACTCAAGA TCAACCATAT GAAAAATGGT TTGAAATTGT AAAAAATCAC TGCAAGATG 720  
 AATCAAATAT TTTAGATATT GGATGCGGTA CTGGTAGTTa ACAGTTCAAT TAGAAGCTTT 780

55

## (2) INFORMATION FOR SEQ ID NO: 371:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1823 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 371:

|             |             |             |            |             |              |      |
|-------------|-------------|-------------|------------|-------------|--------------|------|
| ATAGATGAAG  | GTGCAAAATAT | TGAAaTAGGT  | TATTACCTG  | GACGCTTGAA  | ATGGTTAGTT   | 60   |
| GCTGATTTAT  | TAACTAAACA  | AGGATTAAAA  | GTAGTTAACG | ACGATATGAC  | AGGAAGAACG   | 120  |
| TTAAAAGATC  | GTAATATTAT  | AACAGGTGAC  | AGTCCITTAG | CTTCAAATGA  | GTTAGGAAAA   | 180  |
| TTAGCAGTTA  | ATGAAATGTT  | AAATGCAATA  | CAAAATAAAT | AATTAATAT   | TAATTAGAGG   | 240  |
| AGCCTCATAT  | GTAATGTAT   | GAGGGCTCTT  | TTTTTTGGCA | AAATTAAAGT  | GATACTTGTA   | 300  |
| AAATAGAACC  | TATTATGAGT  | ATGATTTAAG  | AAAACGCTTG | CAAACTAAT   | AACCGCAACT   | 360  |
| AGCGATATGG  | AGGAAACATG  | ATGTCCTTATA | GCATTGGAAT | TGATTATGGA  | ACTGCTTCAG   | 420  |
| GCCGTGTGTT  | TTTAATTAAAT | ACAACCTAACG | GTCAAGTAGT | ATCAAAATTT  | TGAAAACCAT   | 480  |
| ATACACATGG  | TGTCATTGAG  | AGTGAATTAA  | ATGGTTTGAA | AATACCACAT  | ACATATGCAC   | 540  |
| TTCAAAATAG  | TAATGATTAT  | TTAGAAATTA  | TGGAAGAAGG | AATATCATAT  | ATAGTACGTG   | 600  |
| AATCAAAAAT  | AGATCCAGAC  | AATATAGTAG  | GTATTGGTAT | AGACTTTACT  | TCATCTACTA   | 660  |
| TTATTTTTAC  | TGACGAAAAC  | CTTAACCCGG  | TACATAACTT | AAAACAATTT  | AAAAACAATC   | 720  |
| CACATGCGTA  | TGTGAAACTT  | TGGAACATC   | ATGGTGCATA | TAAAGAGGCA  | GAGAAATTAT   | 780  |
| ATCAAACTGC  | TATTGAAAAT  | AATAATAAGT  | GGTTAGGCCA | TTATGGATAT  | AATGTTAGTA   | 840  |
| GTGAATGGAT  | GATTCCTCAA  | ATAATGAGG   | TCATGAATCG | AGCACCAGAA  | ATTATGGAAA   | 900  |
| AAACGGCTTA  | TATTATGGAA  | CGGGGCGATT  | GGATTGTAAA | TAAATTAACT  | AATAAAAAATG  | 960  |
| TACGCTCGAA  | TTGTGGATTA  | GGTTTCAAAG  | CATTTTGGGA | AGAAGAAAACA | GGGTTTCATT   | 1020 |
| ATGATTATTAT | TGATAAAATA  | GACCCCAAAT  | TATCAAAGT  | AATTCAGAT   | AAAGTATCTG   | 1080 |
| CACCGGTGTT  | TAATATTGGT  | GAAGCAGTAG  | GGAAACTGGA | TGATAAAATG  | GCACAGAAAT   | 1140 |
| TAGGATTATC  | AAAAGAACT   | ATGGTAAGTC  | CTTTTATTAT | TGATGCCCAT  | GCTAGTTTAT   | 1200 |
| TAGGTATTGG  | GTCGTAAAAA  | GATAAAGAAA  | TGACTATGGT | GATGGGAACA  | AGCACATGCC   | 1260 |
| ATCTTATGTT  | AAATGAAAAA  | CAACATCAAG  | TGCCAGGTAT | ATCAGGTCTC  | GTA AAAAGGAG | 1320 |
| CAATTATTCC  | AGAATTATTT  | GCTTATGAAG  | CGGGGCAATC | AGCAGTAGGT  | GATTTGTTTG   | 1380 |

CTGTATTGA ATTAATGAAT GAAAAGATAA AACATCAAT GCCAGGTGAA AGTGGGCTCA 1500  
 TTGCTCTTGA TTGGCATAAT GGAAATCGAA GTGTATTAAG TGATAGCAAT TTAACAGGTT 1560  
 5 GTATCTTTGG ATTAACTTTA CAACTAAGC ATGAGGATAT TTATAGAGCm TATTTaGaaG 1620  
 CTACAGCATT TGGTACTAAG ATGATTATGC mACAGTATCA AGATTGGCAT ATGGaAGTAG 1680  
 aAAAGGtATT TGCaTGTGc gGTaACCTa AAAAGAATGC TGTTATGATG GATATCTATG 1740  
 10 CGAATGTACT GAATAAAAAA CTAATTGTTA TGGATAGTGA GTATGCACCA GCAATAGGCG 1800  
 CAGCAATATT AGGTGCAGTC AGT 1823

## (2) INFORMATION FOR SEQ ID NO: 372:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1600 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 372:

ACGATCATCT GCATCAGCGA ATTCCGATGc AaTTcATATG tCCtAATAAC AAAAGAATAG 60  
 GTTTTAAAG ATATGCATAC CGTAATGaTA GATATAGTTT TAAACGTGAC TTCAAGCTAT 120  
 ATGaATGTGA TGACTGTTCA TCATGTTCTT TGAGACATCA ATGCATGAAG CCAAAATCGA 180  
 30 AATCCAATAA GAAAATTATG AAGAATTATA ATTGGGAATA CTTTAAAGCC CAAATTAAATC 240  
 AAAAGCTTTC TGAACCGAA ACGAAAAAAT CTATAGTCAA AGAAAAATTG ATGTAGAGCC 300  
 TGTTTTTGA TTTATGAAGG CTATTTTGGG TtTCACTCGA ATGTCAGTTC GGGAAATAAT 360  
 35 AAAGTTAAAC GAGAGCTAGG TTTGTATTA ATGGCACTTA ATATAAGGAA AATAgcagCT 420  
 CAACGAGCTG TACATTATAA AATACATATC AAAAAAGCTG ATTTCTATCA AATAAATAAT 480  
 AGAAATCAGC TTTTTACAT TGCCTAAGAA CTTTAAGGAA CTTAATGTCC CAAGCTCTTT 540  
 TTTGTTATAT CTAATTGTTA ATTTATGATT GTTTATTCGG TCCTTTGATG TTCACTAAAT 600  
 GTGACTTTAA ATCTTGTTCT AATTGTTGTA ATTCTTTTT AGCTAATTGT CGTTCCTCGC 660  
 45 GACCGTGTG TTGAATAATT AATGTTTCTT CAATTGTCTC AATAATGTTA CGCTGTGTAC 720  
 GTTTCAATGT ATCAAGATCA ACAATGCCAC GCTCAATTTT TGTTCAGTT TCAATCGCAT 780  
 TTGTTTCAA CAATTTCAGCA TTTGCTGTTA ATAAATCAAT AGTTGTATCA GTGACAGCTC 840  
 50 GTTGTGCAGC AACTGCATTA CGCTGTCTCA TTAATGTAAG CGCAATGGCC ATTTGATTTT 900  
 TCCATAGTGG AATACTTGTC AAAATTGAAC TTTGTATCTT CTCGGCAAGT GCTTGATTAA 960

AGTCATATAT GCGTTTATCT AGTCTATCTA TAAATTGCTG CATATCTGCA ACTTGTGAA 1080  
 TATCCATTG ATTAGTGGAT TgtGCGCTTG CTGTTGCAAT TGTGGTAGCT TTTCATTTTC 1140  
 5 TAATTGCAAC TTTTTTGTcT GTGCAGCAAT GATATGCAAT GATAAGTCAT CAAAGTATTG 1200  
 TTGTGTTTA TCATATAGCG TATCTAATAA TTCAATATCT CTGTGTAAT GTGTTTGATG 1260  
 TTCTGTCAGT TGAATCGTTA TGGATCGAC TTGAGCACTA ACTGATTGCA TTCTTGAAAA 1320  
 10 GATTCATTG ATAGACGACT TTGCTCTGCT AAAAATICTT TTAAACATAG ATGGTTTATC 1380  
 AGTATTTAAC TCATTGGAT TAACGTACTT TAGTTTGAC ATTAGATCTG ACAAAGTATC 1440  
 TCCAATAGGA CCAACATCTT TACTTTGTAC TTCATCCAAC ATTTGATGTG AAAATTGAGA 1500  
 15 CATTTGTTTC TGKAAATCAG AACCAAAACG TAATAAACCT TCATTGTCTA AAGGGTTAAT 1560  
 TTGTTTACTG ATTGTGTCTA CCTGTTTTG TTGTCAATT 1600

(2) INFORMATION FOR SEQ ID NO: 373:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1227 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 373:

ATGTTGATAA TGTTAGCTTC TACAAGAATA AAGACCAACA AGTTGCTGCA ACAATTCTTG 60  
 aTAGTAAAC TGTTGGTTTA GTTGCTATAT CTGGTGGACG TGATTTCAAA GACGTCGTTA 120  
 ACAGAAACCA AGCAACAGAT CCTCACCCCTA CTGGTTCATC TTTAAACCT TTCTTAGCGT 180  
 35 ATGGACCTGC CATTGAAAA ATGAAATGGG CAACAAACCA TGCGATTCAA GATGAATCTT 240  
 CATATCAAGT TGATGGTTCT ACATTAGAA ACTATGATAC GAAGAGTCAC GGTACTGTAT 300  
 CTATTATGA TGCTTTACGA CAAAGTTTCA ATATCCACG TTTAAAGCT TGCGAATCAG 360  
 40 TTAAGCAAAA TGCTGGTAAT GATGCACCTA AGAAATTCGC TGCCAAACTT GGCTTAAACT 420  
 ACGAAGGCGA TATTGGTCCA TCTGAAGTAC TTGGTGGTTC TGCTTCAGAA TTCTCACCAG 480  
 45 CACAATTAGC ATCAGCATTG GCTGCAATCG CTAACGGTGG TACTTATAAC AAGCGCATT 540  
 CAATTCAAAA AGTAGTTACT CGTGATGGTG AAACAATCGA ATACGATCAT ACTAGCCATA 600  
 AAGCGATGAG TGATTACACT GCATACATGT TAGCTGAGAT GCTAAAGGT ACATTIAAAC 660  
 50 CATATGGTTC TGCATATGGC CATGGTGTAT CTGGAGTAAA TATGGGTGCT AAGACAGGTA 720  
 CTGGTACTTA CGTGCTGAA ACTTATTCAC AATATAATTT ACCTGATAAT GCAGCGAAAG 780

EP 0 786 519 A2

|    |            |             |            |            |            |            |      |
|----|------------|-------------|------------|------------|------------|------------|------|
| 5  | AAGTTAAACA | ATATGGTGAA  | AACTCATTTG | TGGACATAG  | CCAACAGAA  | TATCCACAGT | 900  |
|    | TCTTATATGA | AAATGGTGAT  | TCAAAAATT  | CATCTAGAGA | TGGCGAAGAC | TTTAAAGGTC | 960  |
|    | CTAGCTCAGT | AAGTGGTAGT  | ATCCCATCAA | TCAATGTTTC | TGGTAGTCAA | GATAACAACA | 1020 |
|    | CTACAATCG  | TAGTACACAC  | GGTGGTAGTG | ACACATCAGC | AAACAGCAGT | GGTACTGCAC | 1080 |
| 10 | AATCAAATAA | CAATACTAGA  | TCTCAACAAT | CTAGAAACAG | CGGTGGATTA | ACAGGTATAT | 1140 |
|    | TCAACATAC  | CACCTCAACAT | AAAATCTCTA | GTTATACCAT | ATTTATGGTG | TAGCCGAGGA | 1200 |
|    | TTTtTtTAGG | TTCTTCATCT  | TTTATGG    |            |            |            | 1227 |

(2) INFORMATION FOR SEQ ID NO: 374:

(i) SEQUENCE CHARACTERISTICS:

- SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1953 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 374:

|    |             |              |             |             |             |             |     |
|----|-------------|--------------|-------------|-------------|-------------|-------------|-----|
| 25 | CCATATTGGtG | CACATCTCTT   | CGTATTTAGT  | GATTATTTAA  | AACCAAGCGT  | ACGTTTATCA  | 60  |
|    | TCAATTATGG  | GATTAATAATgC | acAGCTTCACT | TTCCACACATg | AyTcaATTGc  | AGTAGGTGAA  | 120 |
|    | GATGGTCCTA  | CTCATGAACC   | AATTGAGCAA  | TTAGCTGGAT  | TAAAGAGCCAT | TCCAAATATG  | 180 |
| 30 | AATGTTATCC  | GTCTCGTCTGA  | TGGTAATGAAC | ACAAGAGTAGT | CATGGGAAGT  | TGCGTTTAGAA | 240 |
|    | TCGTGAATCA  | CACCTACTTC   | ATTAGTATTG  | ACACGTCAAA  | ACTTACCGGT  | ATTAGATGTG  | 300 |
|    | CCGAGAAGATg | TAGTGTGAAGA  | AGGCGTTCGA  | ACAGGTGCTC  | ATACAGTTTA  | TGGCTCTGAA  | 360 |
| 35 | GAGACACCAG  | AATTCCTATT   | ATTAGCTTCA  | GGTTCAGAAG  | TTAGTCTTGC  | AGTTGAAGT   | 420 |
|    | GCTAAAGATC  | TTGAAAAACA   | AGGTAATCA   | GTACGTGTTG  | TTTCAATGCC  | TAACTGGAAT  | 480 |
|    | GCATTTGAAC  | AAACAATCTGA  | AGAATATAAA  | GAATCAGTTA  | TTCCATCAAG  | CGTAACAAAA  | 540 |
| 40 | CGTGTTGCGA  | TTGAAATGGC   | TTACCGCTT   | GGATGGCATA  | AATATGTAGG  | TACTGCGAGT  | 600 |
|    | AAAGTTATTG  | CTATTGACGG   | CTTTGGCGCA  | AGTGCACTG   | GCGATTTAGT  | AGTTGAAAAA  | 660 |
| 45 | TATGGATTTA  | CAAAAGAAAA   | TATCTTAAAC  | CAAGTTATGA  | GCTTTATAAG  | ATAATTTATA  | 720 |
|    | AAGCGAGTAT  | GTTTAGAAGT   | CTAGGATGCA  | TAACTTAGG   | CTTCTTTTAA  | AGTGTGTAAA  | 780 |
|    | TTTAGAGTAT  | AGCACTTAA    | CTCATCATCA  | AGTGATAAGT  | TATGAAAAGT  | TACTATTTC   | 840 |
| 50 | GATTAAGCTT  | TAAAGGCTCT   | GTATATAACG  | CATGATTTTT  | GATATATT    | TTAGTATCGA  | 900 |
|    | TATTAATAAT  | CTTGAATATA   | CTGATTTCTG  | ATAAATGTG   | ATGAATTTAG  | TAAAAATCG   | 960 |

TGCATTAATT TTAGGTTTAA TTGGAGGTTT CCTTTTAGCT AGAAAATATA TGATGGACTA 1080  
 CTGGAAGAAA AACCCACCAA TCAACGAAGA AATGCTTCGT ATGATGATGA TGCAATATGGG 1140  
 5 TCAAAAACCT TCTCAGAAGA AAATTAATCA AATGATGACG ATGATGAATA AAAATATGGA 1200  
 TCAAAATATG AAGAGTCCGA AAAAGTAAAT TCGCAATTGA TAGAGGCTAT TTTCCAGATA 1260  
 TGGAAATGGC CTCCTTTTAT AATCAAAATTA ATAAGAAATA ATATGTTTAT TAAAATTAAA 1320  
 10 GTTAACAAAA TGACGAATAG ACTGAGAAAT GCTATAATTC ATTTTGATATG ATTTACAGAG 1380  
 AGTTTATTTA ACGAGAAGGT GTCyGCGTGC TCTATTTAAT ATTTCAACT ATTGTAGCTT 1440  
 TATTTATGGG AACTATAGTT ATAGTTATTC GTATGAAAGC TCAAAATTAT CCGGTAAAKG 1500  
 15 AGAAAAAAT AGTTTTGCCA CCGTTTTTAA TGGCGACCGG TGCATTGATG TACGTCGTTT 1560  
 CAtATTTTAT GCTAACAGGA TCGGAAATGC TAGAAGCCTT TATAATTGGT TTGCTTTTTT 1620  
 20 CcACAGTTCT AATTITGGACT TCTCGATTG AAGTCAAAGG TACAGAAAT TATATGA AAC 1680  
 GATCTAAAGC ATTCCAGTT ATTTTGATTT CATTACTTAT CATTGCTACT GTGATGAAAA 1740  
 TATTCTTAGT TAATGAAATA GATCCTGGAG AATTAGGCGG CATGTTCTTT TTATTAGCAT 1800  
 25 TCTGTATGAT TGTTCCCTGG AGAGCAGCAA TGCTATATAA ATACAAAAAA CTAAAGAAAA 1860  
 CATTAAATCAA TTAATTACTT TtAAACCAC TTGTGATCGA CTCTTAAATC AGTCAATGAG 1920  
 TGGGTTTAAT nTTACTTGGA AAAGnGGAA AGG 1953

(2) INFORMATION FOR SEQ ID NO: 375:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 3787 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 375:

ACATTGTGAT AAntTATCGA CATTAAAGAT GAATTCAnTT GATCGTnTCA ATGATTATCC 60  
 TGTGGAAGTA GCACGTTTGC TTGATATAGT GGAmATaAAA GTACACGCAT TACATTACAGG 120  
 45 TATCCACGTT GATTAAAGAT AAAGGGAAAA TAATTGATAT TCATTTATCT GTAAAAACCA 180  
 CTGAAAAATAT TGATGGCGAA GTGCTGTTCA AAGCAACACA ACCTTTAGGT AGAACAAATGA 240  
 AGGTTGTGTG TCAAAATAAT GCAATGcCAA TTACTTTAAC GAAACAAAAAT CAATGGCTTG 300  
 50 ATAGTTTGAA GTTTTTAGTT AAGTGCATTG AAGAAAGTAT GAGAATCAGT GATGAAGCAT 360  
 AAAGAAGCAT TTAATGGCGT TGTCGTGTTA ACTGCTGCAT TAATTGTCAT TAAAAATCTG 420





EP 0 786 519 A2

|    |             |             |             |            |             |             |      |
|----|-------------|-------------|-------------|------------|-------------|-------------|------|
| 1  | ATGATGGCTT  | CACACTGTTA  | GATCGGCACAT | CAITACAAGA | AGTAACACTT  | AATGTTAGAA  | 2340 |
| 2  | CGCATACATT  | GATTACGCCAA | GTTTATAGTG  | CAATGGTTGC | TGCTAAITTTG | AAAATCACTT  | 2400 |
| 5  | TAATGGAAACG | ATATCCTGAT  | GATTACCCCTG | TTCAAATTGT | CACCTGGTGCA | CGAAGCGATG  | 2460 |
| 6  | GTGCGGATAA  | CGTTGTGACA  | TGCCCAATTAT | ATGAATTGGA | TCATGATGAA  | AATGCATTCA  | 2520 |
| 7  | ATAATTTGAC  | GAGTGTATTTC | GTACCAAAAA  | TCATAACATC | GACATATTTG  | TATCATGACT  | 2580 |
| 10 | TTGATTTTGC  | AACGGAAAGT  | ATTGATACCTT | TAGTTGATGA | AGATAAAGGT  | TGTCCATGGG  | 2640 |
| 11 | ATAAAGTGCA  | AACGCaTgmA  | AcgCTAAAGC  | GTTATTTACT | TGAAGAAAAA  | TTTGAATTGT  | 2700 |
| 12 | TCGAAGCTAT  | TGACAATGAA  | GATGATTGGC  | ATATGATTGA | AGAACTAGGA  | GATATTTTAT  | 2760 |
| 15 | TACAAGTGTT  | ATTGCAATCT  | AGATTTGGTA  | AAAAAGAAGG | GATATCGAC   | ATTTAAAGAAG | 2820 |
| 16 | TGATTACAAG  | TCCTTAACTG  | AAGTAATGTC  | TAAGACACCC | ACACATATTT  | GGTGAATGCCA | 2880 |
| 20 | ATGCTGAAAC  | TATCGATGAC  | TTAAAGAATA  | TTTGCGCTAA | GCGCAAGAAGT | CGTGAAGGTA  | 2940 |
| 21 | AACAGCCAAG  | AGTTAAATTT  | GA AAAAGTAT | TTGCAGAGCA | TTTTTTAAAT  | TTATATGAGA  | 3000 |
| 22 | AGACGAAGGA  | TAAGTCATTT  | GATGAGGCCG  | CGTTAAAGCA | GTGGCTAGAA  | AAAGGGGAGA  | 3060 |
| 25 | GTAATACATG  | AGATTAGATA  | AATATTTAAA  | AGTATCACGG | TTAATAAAGC  | GACGTACGCT  | 3120 |
| 26 | AGCAAAAGAA  | GTAAGTGATC  | AAGGTAGAAT  | TACAATAAAT | GGTAATGTTG  | CTAAAGCTGG  | 3180 |
| 27 | ATCGGATGTT  | AAAGTTGAAG  | ATGTGCTGAC  | GATTGCGCTT | GGTCAAAAAT  | TAGTAACAGT  | 3240 |
| 30 | TAAAGTAACT  | GCATTAAATG  | AACATGCATC  | TAAAGATAAC | GCGAAGGGTA  | TGTATGAAAT  | 3300 |
| 31 | CATTGAAGAG  | CGTCGACTTG  | AAGAAGCGTA  | AATTGGAGGT | GACAAGCAAT  | GA AAAATAAA | 3360 |
| 32 | GTAGAACATA  | TAGAAAAATCA | GTACACGTGC  | CAAGAGAACA | AGAAAAAACA  | ACGTCAAAAA  | 3420 |
| 35 | ATGAAAAATGC | GTGTTGTTG   | TAGGCGTATT  | ACAGTATTTG | cGGGCGTATT  | aCTTGCGATA  | 3480 |
| 36 | ATTGTTGTTT  | TATCaATCTT  | GCTTGTGTGC  | CAAAAACATC | GCAATGATAT  | TGATGCACAG  | 3540 |
| 37 | GAGCGAAAAA  | CGAAAGAAGC  | ACAGTTTCAA  | AAGCAACAAA | ATGAAGAAAT  | TGCGTTAAAA  | 3600 |
| 40 | GA AAAGTTGA | ATAATCTGAA  | TGACAAAGAT  | TACATTGAAA | AAATTTGCCG  | TGATGATTAT  | 3660 |
| 41 | TACTTAAGCA  | ACAAAGGTGA  | AGTGATTTTT  | AGGTTGCCAC | AAGACAAGA   | TTCTGCTAGC  | 3720 |
| 45 | TCAAAATCTT  | CGAAAAATAA  | AATCCAAATT  | GATTCCAAAT | TATCCGAGTA  | TAGACATTGT  | 3780 |
| 46 | GA AAAAAA   |             |             |            |             |             | 3789 |

(2) INFORMATION FOR SEO ID NO: 376:

50 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1644 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double

EP 0 785 519 A2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 376:

|    |   |      |
|----|---|------|
| 5  | TAAACCATTT CAACTGAGG AACGCAAGA CGGACgTTT CCAGATTTAG AAGTATTTAA      | 60   |
|    | AAATGAATGT GATTTAAGCT ATGACATAAC GTCACCTTAT ACTTTTAAGC AACCTGTATC   | 120  |
|    | ACCACACCTT GCATTTAAAA TGACAGATCA AATTTTCTA AATAAGCAGC GTGTATTAGA    | 180  |
| 10 | TAAGTAAAA GTTTTAGATA AGGAATTTGA TTTTATCTTA ATTGAGGGTG CTGGGGGAAT    | 240  |
|    | TGCCGTACCA ATATATGAAG GTACAGATGA TTTCTACATG ACTAAAGATC TAATCAATGA   | 300  |
|    | TTGTGCAGAT TGTGTCATCA GTGTGTTGCC ATCAAAATTA GGTGCTATTA GCGATGCCAT   | 360  |
| 15 | TGTTTACCAA GATTATGTTA ATCAGAATGT ATCGGCGAGT AATTTTTTAA TAATGAATCG   | 420  |
|    | CTATACAGAC AGCTATATTG AAAAAGACAA TCAATGACG ATTGGAATTA TAACAAATAA    | 480  |
| 20 | AACAGTCTAT ACATTTGAAG AACATGCCAC GTATGAAAT TTCTCAGAAG CATTTTAAAA    | 540  |
|    | ACAATTAAAT GGAGTTAAAA ATGAATTACA CACAACACT TAAACAAAA GACTCAGAAT     | 600  |
|    | ATGTTTGGCA TCCATTTACA CAAATGGGTG TATATAGCAA AGAAGAAAGC ATCATCATTG   | 660  |
| 25 | AAAAAGGAAA GGGTAGTTAC CTTTACGATA CGAATGGCAa TAAATATTTA GATGGTATG    | 720  |
|    | CATCGTTGTG GGTCAATGTG CATGGTCATA ATAACAAATA CTTTAAATAG GTAATTAAAA   | 780  |
|    | AGCAACTCAA TAAATTGCC CATTCTACGC TGCTAGGATC ATCAAATATT CCGTCAATAG    | 840  |
| 30 | AACCTGCGGA AAAATTAATC GAAATCACGC CAAGTAATCT AAGAAAAAGTA TTTTATCTG   | 900  |
|    | ATACAGGCGAG TGCCTCTGTT GAAATCGCAA TAAAGATGGC ATATCAGTAT TGGAAAAATA  | 960  |
|    | TTGATAGAGA AAAATATGCC AAGAAAAACA AGTTTATAAC GCTAAATCAC GGTATCATG    | 1020 |
| 35 | GGGATACGAT TGGTGCGGTA AGTGTGGTG GTATCAAGAC CTTTCATAAA ATATTTAAAG    | 1080 |
|    | ACTTAATATT TGAGAATATT CAAGTAGAAA GCCCATCTT CTATCGCAGT AATTACGATA    | 1140 |
| 40 | CTGAAAATGA AATGATGACA GCTATTTTAA CGAATATAGA GCAAAATCTA ATTGAAAGAA   | 1200 |
|    | ATGATGAAAT CGCAGGGTTT ATATTGGAAC CGTTGATTCA AGGTGCGACA GGCTTGTTTG   | 1260 |
|    | TTTATCTCTAA AGGCTTTTTG AAAGAAGTCG AGAAATTGTG CAAAAATATC GATGTCCTTAT | 1320 |
| 45 | TAATTTGTGA TGAGGTAGCA GTTGGTTTTG CGAGAACTGG AAAGATGTTT GCATGCAATC   | 1380 |
|    | ATGAAGATGT TCAACCGGAT ATTATGTGTT TAGTAAAGG GATTACTGTT GGCTACTTAC    | 1440 |
|    | CACITGCGAG TACATTGACA TCTAAAAAAA TATACAAATG ATTTTAAAGT GATTGCGATG   | 1500 |
| 50 | GTGTGAATAC CTTTTCATC GGTcATACAT ACaCCGGAAA TCAAAATcGTT TGTaCGTGTG   | 1560 |
|    | cATTaGaAAA TATaAGaCTT TATGaAAAAA GTaAGTtnAT TGTgCACATa TTGaAACGaC   | 1620 |

55

## (2) INFORMATION FOR SEQ ID NO: 377:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 431 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 377:

ATGATTTTtTA aAAATCATTa AGTTaAGGTr GATACACATC TTGTcATATG ATCAaATGGT 60  
 TTGcCCAAAA ATCAATAATC AGACAACAAA ATGTGCGAAC TCGATATTTT ACACGACTCT 120  
 CTTTACCAAT TCTGCCCCGA ATTACACTTA AAACGACTCA ACAGCTTAAC GTTGGCTTGC 180  
 CACGcTTTAC TTGACTGTAA AACTCTCACT CTTACCGAAC TTGGCCGTAA CCTGCCAAC 240  
 AAAGCGAGAA CAAAACATAA CATCAACGA ATCGACCGAT TGTAGGTAA TCGTCACCTC 300  
 CACAAGAGC GACTCGCTGT ATACCGTTGG CATGCTAGCT TTATCTGTTT GGGCAATACG 360  
 aTGCCcATTG TACTTGITGA CTGGTCTGat ATCcGTGAGC AAAAaCGGCT TATGtTnTTG 420  
 CGAGCTTCAG T 431

## (2) INFORMATION FOR SEQ ID NO: 378:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 2006 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 378:

TTTnTATAAC GTATtATRAA TCGtTAAAAA TTTTGgTTGT GTTTGCGTCA CGTAGACAAC 60  
 CTCcATAAAG TTACTTAATC ACTCTCATCA TACAATAATT TTTACTCAAA TTGGAAaAAT 120  
 TATAAAAAATT AAATATAGAT AGGCTTTGAA AATTAGTTTT ATACAAGGTT AGTAGCTGTA 180  
 ACTGTAAAAAT GTTCTTAATA TTGTCAAAAT GTAATGCTTG AAAGCGCTTT TAAaAAATAT 240  
 TATTATATAC ATGGTTAGAC AAATAGACAA ATCACTATAC AAATATTGGG AGGAATATTT 300  
 TATGAaATCA ACACCACACA TTAAACCAAT GAATGACGTC GAAATTGCAG AAACGGTTCT 360  
 ATTGCCAGGA GATCCGTTAA GAGCTAAGTT CATTCGAGAA ACTTATTGG ATGATGTGGA 420  
 ACAGTTCAAT ACAGTGCgAA ACATGTTTGG TTTTACCGGA ACATATAAAG GTAAAAAAGT 480  
 TTCTGTcATG GGTTCAGGTa TGkGTATGCC ATCTATTGGC ATTACTCTT ATGAATTAAT 540

CATTGATTTA TATGATGTGA TTaTTkCACA AGGTGCCTCT ACTGATTCAA ATTACGTTCA 660  
 ACaATATCAA TTACCAGGTC ATTTTGGCGCC AATTGCTTCT TATCAATTAT TAGAAAAAGC 720  
 5 AGTTGAAACA GCACGTGACA AAGGTGTAGC TCATCATGTA GGTAATGTGT TATCAAGTGA 780  
 TATTTTCTAT AACGCGGATA CAACAGCGAG TGAACGTTGG ATGCGTATGG GTATTTTAGG 840  
 TGTAGAAATG GAATCaGCTG CaTTATACAT GaATGCAaTT TACGCTGGTG TCGAAGCATT 900  
 10 AGGTGTGTTT ACAGTGAGCG ATCaTTTAAT TCATGAAACG TCAACAAcAC CTGAGGAAAG 960  
 GGAACGTGCA TTTaCAGATA TGATTGAAAT TGCACGTGCA TTGGTGTAGA TGATTATGAA 1020  
 TGTGTAATAT TCTAAAATAA AGAAAGCAGT ACCTATTTTA TTATTCTTAT TTGATTTCAG 1080  
 15 TTGGTTATA GACAACTCAT TTAaATTGAT TTCTGTAGCC ATTGCTGATG ACTTAAACAT 1140  
 ATCTGAACG ACAGTAAGTT GGCAAGCGAC ATTAGCCGGT TTAGTAATTG GTATTGGCGC 1200  
 TGTAGTATAC GCTTCATTAT CTGATGCCAT TAGTATACGC ACACATTTTA TTTATGGCGT 1260  
 GATATTAACT ATTATCGGAT CAATTATTGG TTACATTTTC CAACATCAAT TCCCACTACT 1320  
 TTTAGTTGGA CGTATTATTC AAACGCGCG TTTAGCTGCT GCAGAGACAT TATATGTGAT 1380  
 25 ATATGTTGCA AAGTATCTTT CTAAAGAGGA CCAGAAGACT TACCTTGGCT TAAGTACGAG 1440  
 CAGTTATTCC TTGTCAATAG TTATCGGTAC ATTATCAGGT GGATTIATIT CTACGTATT 1500  
 ACACGTGACA AATATGTTTT TAAITGCATT AATCGTAGTA TTTACGTTGC CATTCTCTATT 1560  
 30 TAAATTATTA CCAAAAGAAA ATAATACGAA TAAAGTCAT TTAGATTTTG TTGGCTTAAT 1620  
 TCTAGTGGCA ACTATTGCTA CAACAGTCAT GCTGTTTATT ACGAaCTTTA ATTGGTTATA 1680  
 TATGATTGGT GCCTTAATTG CGATTATCGT TTTTGGGCTA TATATTAAAA ATGCGCAACG 1740  
 35 TCCATTAGTA AATAAATCAT TTTTCCAAAA TAAACGTTAT GCTTCATTTT TATTIATAGT 1800  
 ATTTGTAATG TATGCTATCC AATTGGGTTA TATTTTTACG TTCCCAITCA TAATGGAGCA 1860  
 AATTATCAT CTGcAACTAG ACACAATC ACTGTTATTA GTACCGGGt TaTATAGTAG 1920  
 40 CAGTCATTGT TGStGgCACT AAGTGgGtTA AAATCGGgCG rAATATCTGG AATTCCAAAA 1980  
 CCAAGCGGAT TATCACAGCC AATTAA 2006

## (2) INFORMATION FOR SEQ ID NO: 379:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 4858 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## EP 0 786 519 A2

|    |             |             |             |             |             |             |      |
|----|-------------|-------------|-------------|-------------|-------------|-------------|------|
|    | TGGGGAAAAA  | AAGACCACAG  | AGTATTAGAA  | GAATCGTTAA  | ATATTTCTAT  | TGAAGAAATG  | 60   |
|    | AATCGTATCA  | TAAAATTAGT  | CGAAGAATTA  | CTTGAATTGA  | CTAAGGAGGA  | TGTAATGAC   | 120  |
| 5  | ATTCTCTCTG  | AAGCACACAG  | CGTGCAATAT  | AATGATGAAA  | TTGCGTCGCG  | AATACACTCA  | 180  |
|    | TTAAAACAAT  | TGCATCCTGA  | TTATCAATTT  | GATACGGGAT  | TGACATCTAA  | AAATCTAGAA  | 240  |
| 10 | ATTAAATGTA  | AACCTCATCA  | ATTGCAACAA  | TTATTTTTAA  | TCTTTATTGA  | TAATGCAATC  | 300  |
|    | AAATATGATG  | TGAAGAATAA  | GAAAAATAAA  | GTTAAGACAA  | GGTTAAAAAA  | TAAGCAAAAA  | 360  |
|    | ATAATTGAAA  | TTACAGATCA  | TGGAATTGGT  | ATTCAGAGG   | AAGATCAAGA  | TTTCATTTTT  | 420  |
| 15 | GATCGCTTTT  | ATCGAGTGGA  | TAAATCTCGT  | TCAAGAAGTC  | AAGGCGGTAA  | TGGACTCGGA  | 480  |
|    | TTATCTATTG  | CTCAAAAAAT  | CATTCAATTA  | AACGGAGGAT  | CGATTAAAAAT | TAAAAGTGAA  | 540  |
|    | ATTAAACAAG  | GAACAACGTT  | TAAAATCATA  | TTTTAATCAT  | GACTGAGACG  | TCAATCAAAG  | 600  |
| 20 | TCATAGGATC  | AATTTTTTAA  | GTACACATTA  | GCTGTGACTA  | ATGTATAAGA  | ACAACTATAA  | 660  |
|    | AACAATAATA  | CAGTGGTTCT  | TTATCATCTA  | TGTTGTACTC  | CCAAAAATTA  | CAATAAAAATA | 720  |
|    | CATCTATAAA  | CCTAGAAGAA  | TCAACGCTTT  | TGTTGATTCT  | TCTTTTATAG  | CAGATAAATAG | 780  |
| 25 | GTAAATCTAC  | TTTAACAAAT  | AACTAAATAG  | TGATATTATT  | ACATTGTAAG  | CGTTTCAACA  | 840  |
|    | TTTTTGTGGA  | GGGTGTAAAA  | TGACTAACGA  | AAGAAAAAGAA | GTTTCAGAGG  | CTCCTGTAAA  | 900  |
| 30 | CTTCGGTGGC  | AATTTAGGTC  | TAATGTTAGA  | TCTATATGAT  | GACTTTTTAC  | AAGATCCATC  | 960  |
|    | ATCTGTACCA  | GAAGATTTAC  | AAGTCTTATT  | CAGCACAAAT  | AAGAATGATG  | ACTCAATTGT  | 1020 |
|    | ACCAGCTTTA  | AAAAGTACAA  | GTAGTCAAAA  | TAGCGACGGC  | ACAATTAAGC  | GTGTCATGCG  | 1080 |
| 35 | TTTAATTGAT  | AATATTCCGC  | AATACGGGCA  | TCTTAAAGCC  | GATATTTATC  | CTGTAAATCC  | 1140 |
|    | TCCAAAAGGG  | AAACATGTAC  | CTAAATTAGA  | GATTGAAGAC  | TTTGATTTAG  | ATCAACAGAC  | 1200 |
|    | TTTGGGAAGT  | ATATCAGCAG  | GAATTGTTTC  | AGATCACTTT  | GCGSACATTT  | ATGATAATGC  | 1260 |
| 40 | TTATGAAGCA  | ATTTTAAAGAA | TGGAAAAAACG | TTACAAAGGA  | CCAATTGCAT  | TTGAGTATAC  | 1320 |
|    | ACATATTAAAT | AACAATACCG  | AACGTGGTTG  | GTTAAAAAGA  | AGAATTGAAA  | CGCCATATAA  | 1380 |
|    | AGTAACGTTA  | AATAATAACG  | AAAAAAGGGC  | ACTATTCCAA  | CAATTAGCGT  | ATGTTGAAGG  | 1440 |
| 45 | GTTTGAAAAA  | TATCTTCATA  | AAAACCTCGT  | TGGTGCAAAG  | CGTTTTTCAA  | TTGAAGGGGT  | 1500 |
|    | AGACGCACCT  | GTACCGGATG  | TACAACGTAC  | TATTACGATT  | GCTGCGAAGG  | AAGGTATTAA  | 1560 |
| 50 | AAATATACAA  | ATAGGCATGG  | CTCACCGTGG  | ACGTTTAAAC  | GTTTTAACGC  | ATGCTTTAGA  | 1620 |
|    | AAAACCGTAC  | GAATGATGTA  | TTTCAGAAAT  | TATGCATACA  | GATCCAATGA  | AATCTTACC   | 1680 |
|    | TGAAGATGGT  | AGCTTCGAGT  | TAACTGCTGG  | ATGGACTGGT  | GATGTGAATAT | ATCACCTTGG  | 1740 |

|    |             |            |             |             |             |             |      |
|----|-------------|------------|-------------|-------------|-------------|-------------|------|
|    | AAGTCACTTG  | AAAATTGTTG | CACCTGTTGT  | TGAGGGGCGT  | ACGAGAGCAG  | CACAAGATGA  | 1860 |
|    | TACACAACGA  | GCTGGGGCTC | CGACGACTGA  | TCATCATAAA  | GCAATGCCAA  | TTATTATACA  | 1920 |
| 5  | TGGCGATGCT  | GCTTATCCTG | GTCGAAGGAAT | TAACTTCGAA  | ACAATGAACT  | TAGGAAACTT  | 1980 |
|    | GAAGGCTAT   | TCTACGGGTG | GTTTCATTGCA | TATTATTACT  | AACAATAGAA  | TTGGATTTCAC | 2040 |
|    | TACAGAACCA  | ATTGATGCAC | GTTCAACAAC  | TTATTCTTACA | GATGTGGCCA  | AAGGTTATGA  | 2100 |
| 10 | TGTGCCAATA  | TTCCATGTCA | ATGCAGATGA  | CGTTGAAGCT  | ACTATTGAAG  | CAATTGATAT  | 2160 |
|    | TGCAATGGAA  | TTTAGAAAAG | AGTTTCATAA  | AGACGTCGTT  | ATTGATTTAG  | TAGGTTATCG  | 2220 |
| 15 | TCGTTTCGGA  | CATAACGAAA | TGGATGAACC  | ATCAATTACT  | AATCCaGTTT  | CTTATCAGAA  | 2280 |
|    | TATTGCAAAA  | CATGACTCTG | TTGAATATGT  | GTTTGGTAAA  | AAGCTTGTTA  | ATGAAGGTGT  | 2340 |
|    | CATTTCAGAA  | GATGAAATGC | ATTCATTATT  | AGAACAAGTC  | CAAAAAGAAC  | TAAACAAAGC  | 2400 |
| 20 | TCATGATAAA  | ATTAATAAAG | CTGATAAAAT  | GGATAATCCA  | GATATGGAAA  | AGCTGCGAGA  | 2460 |
|    | TCTTGCAATTA | CCGTTACAAG | CAGACGAACA  | ATCATTTACT  | TTTGATCACT  | TGAAAGAAAT  | 2520 |
|    | AAATGATGCA  | TTGTTAACAT | ATCCGGATGG  | CTTTAACATT  | TTGAAAAAGT  | TAAACAAAGT  | 2580 |
| 25 | TCTTGAGAAG  | CGTCATGAGC | CGTTTAATAA  | AGAAGATGGT  | TTAGTTGATT  | GGGCACAAGC  | 2640 |
|    | AGAACAACTT  | GCATTTCGGA | CAATTTTACA  | AGATGGTACA  | CCGATTTCGT  | TAACTGGTCA  | 2700 |
| 30 | AGATAGTGAA  | CGTGTACAT  | TCAGTCATAG  | GCATGCCCGT  | TTACATGATG  | AGCAAAACAG  | 2760 |
|    | TGAAACATAT  | ACACCTTTAC | ATCATGTTCC  | TGATCAAAAA  | GCGACATTTG  | ATATACACAA  | 2820 |
|    | TTCTCCGCTT  | TCAGAAGCAG | CAGTAGTTGG  | TTTTGAATAC  | GGCTATAATG  | TGGAAAAACA  | 2880 |
| 35 | AAAAAGCTTC  | AATATTGGGG | AAGCACAATA  | TGGTGATTTT  | GCAAAATATGT | CACAAATGAT  | 2940 |
|    | TTTTGACAAC  | TTCTTATTCA | GTCTCGCTC   | AAAATGGGGA  | GAACTTCAG   | GATTAAACATT | 3000 |
|    | ATTCTTACCT  | CATGCATATG | AGGGTCAAGG  | GCCTGAACAT  | TCATCAGCAA  | GATTAGAGCG  | 3060 |
| 40 | ATTTTTACAA  | TTAGTGCTG  | AAAATAATTG  | CACAGTTGTC  | AACTTATCTA  | GTTCAAGTAA  | 3120 |
|    | TTATTTCAC   | TTATTGCGTG | CACAAGCGGC  | TAGTTTAGAT  | TCTGAACAAA  | TGCGACCATT  | 3180 |
|    | GGTTGTTATG  | TCACCAAAAA | GCTTACTGAG  | AAATAAAACA  | GTTGCAAAAC  | CAATTGATGA  | 3240 |
| 45 | ATTTACTTCT  | GGTGGATTTG | AGCCAATTTT  | GACAGAATCA  | TATCAAGCGG  | ATAAGGTTAC  | 3300 |
|    | AAAAGTTATT  | TTGGCAACTG | GTAAAATGTT  | CATTGATTTA  | AAAGAAGCAT  | TAGCTAAAAA  | 3360 |
| 50 | TCCAGACGAA  | TCAGTATTAC | TCGTTGCGAT  | TGAAAGATTG  | TATCCATTCC  | CAGAGGAAGA  | 3420 |
|    | GATTGAAGCA  | TTACTAGCAC | AATTGCCAAA  | CCTTGAAGAA  | GTGTCATGGG  | TACAGAAGAA  | 3480 |
| 55 | ACCTAAAAAT  | CAAGGTGCAT | GGTTATATGT  | CTATCCATAT  | GTTAAAGTGC  | TAGTTGCGAG  | 3540 |

ATCAGTCACA CGGTAGGCAT ATAAAATGAG TCGTTTCTAC AACATTTTITA AACAGTTCAT 60  
TCAATATTAT TTTTATCTAA TAATnATATT GGGAGGATTA TACCTTTATA CACACCATGC 120

|    |  |      |
|----|--|------|
|    | ATGCTATTTA GCTAAAGCTA AAAGACCAGA CACTATGCAT ATTTCAACTG GAAATATGTG  | 240  |
|    | CGCATACTTA GTTGCAATTA TTGCTGTAT GATTGTGTAC CTTAATAAAG CGCATGTAAG   | 300  |
| 5  | TATCATCGGT ATAATTATTG GTTTAATGAT TTCATATGTT GTAGTIATCA TACGTCCTTT  | 360  |
|    | ACTAAAGGTG AGCAAATAAA TTAAGAAAGA GGTGAGATTA TGGATCAGAA ATCCCCGCTC  | 420  |
|    | GTGAGTTGGA ATTTATTCGG TTTTGaTATC GTTTTCAATT TATCAAGTAT ATTGATGATA  | 480  |
| 10 | CTTGTTACGG cGTTCTTGT TTTTCTACTT GcTATCATT GTACGCGTAA TTTGAAAAAA    | 540  |
|    | AGACCAACTG GCAAACAAAA TTTCTGTGAA TGGATTTTTG ATTTCTGTAG GGAATCATT   | 600  |
| 15 | GAAGGTAACA TGGCTTGAA AAAAGGTGGT CAATTCCACT TCTTAGCAGT AACGCTGaTT   | 660  |
|    | CTGTACATTT TTATAGCTAA TATGTTAGGT CTTCCGTTTT CTATAGTAAC GAAAGATCAC  | 720  |
|    | ACATTGTGGT GGAAATCACC GACAGChGAT GCAACAGTGA CTTTAAcGTT GTCTACAAG   | 780  |
| 20 | ATAATACTGT TAACTCACTT TTATGGAAAT AAAATGCGTG GTACGAAACA ATATCTTAAA  | 840  |
|    | GGTTATGTAC AGCCGTTTTG GCCATTGGCA ATTATTAAAT TTTTGAAGA GTTCACTTCA   | 900  |
|    | ACATTAAACG TTGGCTGCG TTTGTACGGT AACATATTTG CAGGTGAGAT ACTATTAAAC   | 960  |
| 25 | TTACTTGTCT GCTTATTCTT TAACGAACCA GCATGGGGTT GGATTATTAG TATCCCAGGA  | 1020 |
|    | TTAATTGTGT GGCAAGCATT TTCAATATTT GTAGGAACAA TCCAAGCATA TATCTTTATT  | 1080 |
|    | ATGCTTTTGA TGGTTTATAT GTACACATAAA GTGGCAGATG AACACTAAAA ATTTCAATAA | 1140 |
| 30 | TTATATACAA TCACAGGAGG AAATTAAATT ATGAATTTAA TCGCAGCAGC AATCGCAATT  | 1200 |
|    | GGTTTATCAG CATTAGGAGC AGGTATCGGT AACGGTTTAA TCGTTTCAAG AACAGTTGAA  | 1260 |
| 35 | GGTGTAGCAC GTCAACCAGA AGCACGTGGT CAATTAAATG GTATCATGTT CATTGGTGTA  | 1320 |
|    | GGTTTAGTGT AGGCATTACC TATCATCGGT GTAGTAATTG CATTATGAC ATTTGCTGGA   | 1380 |
|    | TAATTAAACAG ATAAAAGAGG TCGGGACAAA CGGCATAGGA CATAATTCAT GATGCATATA | 1440 |
| 40 | TAGTAATATC TTTGAACTTT ATTAATAGT TGAGATATGA ACGCACCATG CCTATCGCAT   | 1500 |
|    | AAATTCAGTA GGTCTTAACC TCGTCGTTTT TTTCTATATA ACACTAGCGA TTAATTTTAA  | 1560 |
|    | GAAAGGAGTG TCATGAACCC GTGACTGAAA CAGCTAACTT ATTCTGTTCTT GGTGCAGCTG | 1620 |
| 45 | GAGCGTTTGA GTGGGTACT GTGATTGTAC AGGTCTTAAC TTTTCATCGT TTAaTTGCGT   | 1680 |
|    | TACTTAAAAA GTTCGATG GGTCCATTGA AAGATGTAAT GGATAAACGT GAAAGAGATA    | 1740 |
| 50 | TTAACAGAGA TATCGATGAC GCAGAACAA GCTAAGTTAA TGACAGAGAA CTTGAAGAAG   | 1800 |
|    | AAAATAAACA AAAACTTAAA GAAACACAAG AAGAAGTTCA AAAGATTTTA GAAGATGCTA  | 1860 |
| 55 | AGGTTCAAGC ACGTCAACAG CAAGAACAAA TTATTCaTGA AGCAAAcGTA CGTGCAAAAG  | 1920 |



EP 0 786 519 A2

|            |            |            |            |            |            |      |
|------------|------------|------------|------------|------------|------------|------|
| TTAATAATCa | AGTATCTGAA | CTATCAGTGT | TAATTGCTTC | TAAAGTTCIT | AGAAAAGAAA | 2040 |
| TTTCTGAACA | AGACCAAAAA | GCATTGGTTG | ACAAGTATCT | AAAAGAGGCA | GGCGATAAAT | 2100 |
| AATGGTAAAA | GTAGCTAACA | AGTATgcTAA | AGCATTATTT | GACGTGTCAT | TAGATACAAA | 2160 |
| TAATTTAGAG | ACTATTAATG | AAGAATTAAC | AGTTATAAAT | GAAGCAGTAA | AAGATAAAAT | 2220 |
| TG         |            |            |            |            |            | 2222 |

(2) INFORMATION FOR SEQ ID NO: 381:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1143 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 381:

|            |             |            |             |             |             |      |
|------------|-------------|------------|-------------|-------------|-------------|------|
| TAACAATTCC | TaTATTcATG  | TTTAATACGA | AACACTACAT  | TTACATTGTA  | ATTCACTATC  | 60   |
| TTTTGAAGTA | ATAAAGTGAT  | TTGTTCAATC | GATAGCTCAT  | TGCTTGTGTC  | GATTGTAACA  | 120  |
| ATTATATGCA | AGTTTTcAGG  | ATTAACACCT | AATCTTTGAA  | TGATTTGTTT  | AATAGTATAA  | 180  |
| TAATCCATCC | AATAAAAAIT  | ACTTCCTTTA | ATATAAATGT  | TTTTAGGTTG  | ATACATTGTA  | 240  |
| CTCCTTTTGG | TAGGCTCAAA  | AGGTATATCA | ATCTCGCGCA  | TACTTGAAGA  | ACTTTGATTA  | 300  |
| GTATCATCAA | ATAATTCAAT  | TATATTTTTA | TCAATTTCAA  | CTAATTGGGA  | ATGGTTAATT  | 360  |
| GAATGAGATG | TTGGACTATA  | TCTTTTCTTA | ATTAATTITAG | GAGTGTITACC | ATACGTTTCT  | 420  |
| TTAAAGAGGT | GTATAAAACG  | AGAATAATGA | TTAAACCATT  | GACTACTTGC  | GATTTCCTCA  | 480  |
| ATAGGCTTCT | TAGTAGTTAA  | AATATCAATT | AAGCAATGCT  | CCAGTCTAAT  | ATGATTAAAA  | 540  |
| TATTGAATAA | AATTACTATA  | AGGTGTCGCT | TAAACATGT   | CACTTAAAGC  | TTTGTTTGTA  | 600  |
| ATACTAACTT | GATTAAATGAC | ATCTTTCCTA | TTTATCTTTT  | TATGGTGGTT  | GTTTGTAAAA  | 660  |
| TAATCGTGCA | CTTCTTCGGC  | TACTAAATGA | CGACTACCAT  | CGTATGTATT  | TAATGACATT  | 720  |
| AATTCAACAC | ACATGTAATT  | AATAATCTTA | TCATTAGCAT  | TATAAGACTG  | TTGTTTAATT  | 780  |
| TTGCTGTAGA | TAAATACTTT  | AATCAAGATT | CTAAAAGTGC  | TAGCAACCTC  | ATTGTGTAAT  | 840  |
| ATCTTGCCAC | GCATCATATA  | GTTTGCATTG | ATAAACTTCA  | AAAAATGTCTT | TGCACCTTACT | 900  |
| TCAACAACAC | AACATACACT  | ATCATCATTA | CCATCAATTT  | GATACAAATC  | ATTCAACATA  | 960  |
| ATAATGGTGA | CATCATTTCT  | TTTTACATCA | TATTGTTGTA  | GATTGATTGA  | AACATACACT  | 1020 |
| GTACCTTGTA | GCCATATAT   | AATTTTAAAG | TTGCTTCTCT  | TCACACTACT  | CATTTTCATA  | 1080 |

ATT

1143

(2) INFORMATION FOR SEQ ID NO: 382:

5 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 506 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

10

(xi) SEQUENCE DESCRIPTION: SEO ID NO: 382:

|    |   |     |
|----|---|-----|
| 15 | GGACCACCTT TTTTACCTGT ACGACTAAG CCAACTGTTT TAAGGACACC AACAGCATGC    | 60  |
|    | CACACTTGAA TAATTCTCTG TGATGGTCTA AAGCGAACGC TATAAATTAA TGGATGGAAA   | 120 |
|    | TCATCAACAA AAATATAATC GGCCTTACCA AGTAAATATG GCAATCTAAA CTTGTCTCTC   | 180 |
| 20 | CATTTGGGTC TATCCGTAAT ATTCTCCTTA AATACGGTTT TAATATCATATA ATCAAAATCT | 240 |
|    | ACTTTTGGC GTAGTAACTC ATCATATACA TACTTGAAAT TCCCTGATAA ATTCCGAGCG    | 300 |
|    | GAATCTGATG TGAATAATAT TGTTTTGCTT CTTTAAATAT GTAGTAACTT TGTAAATATTA  | 360 |
| 25 | AAAATAGCTT TAAATAAGAA ACTCTTACTT TCAAAAGAAG CTTTATGaCC TTGTTTATGA   | 420 |
|    | AGCCAGTGTC cACTGTTCGC AATGAACCCCT GaTTTCyCyT GagGtAAGGk GaTTTCnATA  | 480 |
|    | TCAAAATACAA ATTGGTTAAC GTCAGCT                                      | 506 |

30 (2) INFORMATION FOR SEO ID NO: 383:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 421 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 383:

|    |             |            |            |            |             |             |     |
|----|-------------|------------|------------|------------|-------------|-------------|-----|
|    | AGCATCTGCA  | AtGAGTCTCT | AATACAACGA | TACGTTTTTG | ATCITTTAGGT | ACTTTTACTG  | 60  |
|    | TACCATTITC  | ATCTTTTACC | GAAATAGTAT | CTTTAGTTGA | TGATCTTTT   | TTACTTGAAT  | 120 |
| 45 | TATCCGTATT  | ACCACAAGCT | GCAACTAAAA | GTAAGGCAAC | TATTAATCCC  | AATATACTAA  | 180 |
|    | AAGTTTTTAT  | ACCTCTCATC | GTTCCACTCC | TTAATATGTA | TAACITTCAT  | TATTTATTTA  | 240 |
|    | TTGATAACAA  | TTATCATCTC | CAAGTAGCGT | TCAATCTTTT | TTTATTTTCT  | AAATATGTATG | 300 |
| 50 | ACTATATAAT  | TCCTCTAATA | AITATGACTA | CAATAGGCAC | ATTTCCTTAG  | ACAAAACTACT | 360 |
|    | GATAATGATAT | CATTGCTATA | TCATCTTTGC | AAITATACAA | TTGACACCAT  | TTAGCATGAC  | 420 |

55

## (2) INFORMATION FOR SEQ ID NO: 384:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 862 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 384:

```

TGCATTGAT ACTTTAGCAC ATGAAATTGC TAATCGATTA GTTGCAAATG ACAAGAATGA      60
AGCAACTTTG GAAATGACTA ATAAATGGC AACGATTCGT TTTACAGAAC CTACGCTGAT      120
TGCATTAGCA GGGGGTAATG TCAAAGCTTA CACTGAGCAT ATGACTATAT CTCCTATATA      180
ATTGTATTGT TTGATAAAG GCGATGTTT AAAGTTTAGA GAAACAAGTT ATACATCGCG      240
AGTGTATTTA GCTGTGGAG GCGGATTGTA ATTAGATGCA TGGTTAGGAT CTAACCTCAAC      300
CGACTTTAAT GTAAAAATTG GTGGTTTAA AGGTAGAACA TTACAAGATG GCGATGAAAT      360
AAAGCTTAAG AGAGATTATA CAGCTCGTCA TCATAAGTTA TTTGAAAACC TTGCTCACAC      420
GAAACAAACA GATTGGGGTA TTGATGGATA CGCCTTGTC TTAATTATA TGTCTGATGT      480
ATTTCATGTC GTTAAAAATA AAGGTACGGA AGATTTTAAA GAAGATGCCA TTCAAAGATT      540
TGTGAAACAT GATTATAAAG TAACGAGCAA AGCAATCGC ATGGGGATGA TGCTTGGAAG      600
TGAAAAAATC AAAGCTTTT ATGAAGATAT GCCACCGTAT CAGACTGTCA AAAAAGGAAC      660
GATACAAATT AAGCGTGATG GCACACCTAT TATCCTATTA AATGATCAIT ATACGCTAGG      720
TAGTACCCG CAAATCGGTA CAATCGCAAG TTATCATTTA ACGAAATTAG CACAAAAACC      780
GCAAGGATCA CGTTTGAAT TTCAATTTAT AGATATTTTA ACGGCTGAAA AGAACCTTGT      840
TAAGTATAGT AACTGGTTAA AC                                             862

```

## (2) INFORMATION FOR SEQ ID NO: 385:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1027 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 385:

```

AAATGAGTTG ATGATGGCTC CTATTCATC TATCTCTAGG TATGACAGAA SATAAACGGT      60
CTTCTACCTG TCTTACATTA TCTTTCCAAC GATTATATACC TAAGCGTTTA GAAACAGTT      120

```

CAGTTTACG TCCTACACCT GCTAAACTTT CTAATTCCTT ATGIGTTTGT GGTATTTCTC 240  
 CATTAATTTG ATCAATCAAA GATTGACAAA GTTCTTAAT ATTCTTAGCT TTGTTACGAT 300  
 5 ACAGACCGAT AGAAGGAATA TCATTTCATA GTTCTTCATC ACTGACTGCC AAATAATCTT 360  
 CAGGCGTTTT GTATTTTTTA AACAGCTCAG TTGTACTCT ATTACTAGA ACGTCTGTAC 420  
 ATTGCGTGA CAATAATACA GCAATAGTTA ATTGGAACGG ATTATCATGT TTTAATTCAC 480  
 10 ATTCTGCATC CGGAAACATA TTGCTATAA CATCAATCAT TTCTAATGCT TTTTCTTAC 540  
 TTACCATCAA GGTTCTCCCC GTTTAACCAA TCAAAATTAG GTACCGTTTT AACTGTGTGC 600  
 GTCATTTTCG GTTTATTGAA TTTTCTCTT ATTTTCTAG AATCGTCAAT TGTTTTGACA 660  
 15 TTGTTTTTCT TCCAATTAAG TAAATACGA TCTATATATT TAAAGCTAAG TTTATTCAAA 720  
 CTATTGCGCT CGTCTAATGC CGCTTGATA ATTGCAGTAT CGTGTATTAT AACATCAATC 780  
 20 CATTGATTTA ACGTTTCTAT TTCATATGGA GATAACGGCC TTGCAATGT ATCCTCTAAA 840  
 ACTCTAAATA ATGTGTTAAA TTTTCTTTA CTATTAGCT CTTCGTTTC CATACITTTG 900  
 TGCTTCAATA TATGACTTAA TTTTTCGAAA AAAGGATCTA GATTATATA TTGCGKAAAT 960  
 25 CTACCTTCyT CATCyTTTTG aACTKgtAAT tCTAGCAATT CACGTGTATC AAATTTTGGa 1020  
 TACCATT 1027

## (2) INFORMATION FOR SEQ ID NO: 386:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1006 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 386:

AAGGnTTGGA GGAATTAAT nGATGCGAAA TCCcAGaTAA AGTACACaC GCATGGAATG 60  
 TGATGCATTT AATATTTCCA GTAGTTAGTA CGAGTTTGC AAGCTTTAAA TCTATGTATG 120  
 GGGGCATACC AAAAGATTTT ATAGACTACT TATTIATTGA TGAAGCAGGA CAAGCAATAC 180  
 45 CTCAGCAGC TGTGGGAGCA TTATATCGTT CAAAAAAGT TGTAGCTGTA GGTGATCCGA 240  
 TTCAATAGA ACCGGTTGTG ACTTTAGAAA GTCAITTAAT TGATAACATT CGTAAAAAAT 300  
 ATCATGTTCC GGAATATCTA GTTCTAAAG AAGCTTCTGT GCAGTCTGTT GCAGACAAcG 360  
 50 CCAATCAATA TGGTTTTTGG AAATCTGATG CTACTGATAG TAATCAAAAA ACCTGGATAG 420  
 GCATACCTTT ATGGGTGCAC AGACGATGTT TAAACCTAT GTTCACGATa GCTaACCAaa 480

## EP 0 786 519 A2

|    |   |      |
|----|---|------|
|    | GSTATGACGT TAAAGGAAaC GCAGTTCAAA AACAAATTGT GAAAGAGCAT GGTGAAAAaG   | 600  |
|    | TAGTGGGATT ATTAGCTGAT GATTGGATTG AAGCAATTAA GGAAGTAAaA AATGAACCGA   | 660  |
| 5  | GCTCATTTGT AATATCGCCT TTTTCAGCAG TACAGCAACA GATTAAACGT ATGTAAaAGC   | 720  |
|    | AACAACtACC GACTAGAATT GATTATTGAAC GTACAAAAAT TAATCAATGG GTGCATAAAT  | 780  |
|    | CCATTGGTAC TGTTCACTACT TTTCaAGGTA AAGAGGCTCA GAAGGTGTAT TTTGTAAATAG | 840  |
| 10 | GTACTGATAA TACCaaAGAT GGTGCTGTGA ACTGTGTAAT CGAaaAACCA AACTTGTtAA   | 900  |
|    | ACGTTGCAGT GACCAAGAGCT AAGAAAGAGT TTAgtGTAAT GCGCAACATGC AAAGAATACA | 960  |
| 15 | GATGAaACCA TTTtATGAGA CGATTtTTTA AGNaAGAAAT GTAAAA                  | 1006 |

(2) INFORMATION FOR SEO ID NO: 387:

(i) SEQUENCE CHARACTERISTICS:

- SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 662 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 387;

|    |            |             |            |            |             |            |     |
|----|------------|-------------|------------|------------|-------------|------------|-----|
|    | CGTTTCATCA | AGTACTGAT   | CCTAATTTC  | AAAATATCC  | GGTAGACTT   | GAGAAGGC   | 60  |
| 30 | GTAAATAG   | AAAGGCTTT   | AAACCAACT  | CAAAAGATAG | CGTTATATTA  | TCAGCAGATT | 120 |
|    | ATTCTCAAAT | TGAATTGCGT  | GTTATTGAC  | ACATTACACA | AGATGAGATG  | ATGAAAGAAG | 180 |
|    | CATTTATCAA | CGGCGATGAT  | ATTCATACAG | CAACTGCTAT | GAAAGTATTT  | GGTGTAGAAG | 240 |
| 35 | CTGATCAAAG | CGATGATTTTA | ATGCGTCGTC | AAGCAAAAGC | GTTAACTTTT  | GGAATTGTTT | 300 |
|    | ATGGGATAAG | TGATTATGGT  | TTAAGTCAAA | GTTTAGGTAT | TACTCGTAAA  | AAAGCAAAAG | 360 |
|    | CATTCAITGA | TGATTATTTA  | GCTAGTTTCC | CAGGTGTAAA | ACAAATATATG | TCGTATATTG | 420 |
| 40 | TAAAGATGC  | CAAAGCTTTA  | GTTTACGTGG | AAACATTGCT | ACATCGTCGA  | CGCTATATTC | 480 |
|    | CTGATATTAC | GAGTCGTAAC  | TTTAATTTC  | CGGGCTTTG  | TGAACGTACT  | GCTATGAATA | 540 |
|    | CGCAATACA  | GGGCAGTGCT  | CGAGATATCA | TTAACTGGC  | AATGGTTAAA  | TTTGCTCAAA | 600 |
| 45 | AAATGAAAGA | GACAACATAT  | CAAGCTAAAC | TATTATTACA | AGTACACGAT  | GAATTAAATT | 660 |
|    | TT         |             |            |            |             |            | 662 |

(2) INFORMATION FOR SEQ ID NO: 388:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 669 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 388:

|    |            |            |            |            |             |             |     |
|----|------------|------------|------------|------------|-------------|-------------|-----|
| 5  | TCATCCAAAT | TTTGAAAT   | CCACATTTTA | CATATCGTAA | TTTTTTAGGA  | AAC TAGTGAA | 60  |
|    | TATAACAAAT | CCCTCCTCTC | ATTTTAAAA  | TAGATATATC | ACTTCCCAC   | TTT TACTTAA | 120 |
|    | CTAAACTGCA | ACGGTTCCTA | ATACCAAAAT | CCTGCCCTCT | ATTTT TATCA | ATTCAAGCAT  | 180 |
| 10 | ACTTATTGAA | AAATGTAAAC | GTTTCTCTGA | TAATCATCTG | AAGCGCATTT  | ATTTTATATA  | 240 |
|    | CTAACGTTTG | AAATATACTA | CAGAGTAGAC | CAATTAATGC | TCAAATATT   | GAAAAAGAAT  | 300 |
|    | TAAAAAGAA  | GTATTAGAT  | TTACTATCCC | AAAAATTTGA | TACTCCAGAA  | AAACTTGCAA  | 360 |
| 15 | CTGAAATTAT | CAATTTAGAA | TCAATTTTAG | AATTACCTAA | AGGTACGGAA  | CATTTCTGCA  | 420 |
|    | GCGATTACA  | TGGTGAATAT | GAAGCTTTCC | AACAGCTATT | ACGCAACGGT  | TCTGGGAACG  | 480 |
|    | TGGAGCGGAA | AATCAATGAT | ATTTTCAAAG | AGAGACTTTC | AACTAAGGAG  | CTTAATGACT  | 540 |
| 20 | TAACTGCTCT | TGCTACTAT  | CCAGAAGrCm | AATTTAAAA  | TgATTAAGAA  | TGATTCCAA   | 600 |
|    | AATTYCGGtC | maCTTAATGt | CyGGtATATC | AcAaCmATCG | aACATTTAA   | TGAGTTAAAT  | 660 |
| 25 | AAATATTGT  |            |            |            |             |             | 669 |

(2) INFORMATION FOR SEO ID NO: 389:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1249 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEO ID NO: 389:

|    |            |             |            |            |            |            |     |
|----|------------|-------------|------------|------------|------------|------------|-----|
|    | CACATGGCTG | TTAGAGATAT  | GAATGGCCAT | CGGTTACCTT | TAACAAAAGA | TGSCAATTTT | 60  |
| 40 | TATCAAACGA | ATGTAGATGC  | AAATGGTGT  | AATCATGGTG | GTAGTGAAT  | GGTGCAAAAT | 120 |
|    | AAACAGGTC  | ATATGAGTCA  | ACAAAGCCAT | ATGAATCAGA | ACACACATGA | ACCAACAGCC | 180 |
|    | ACACATGCAA | CAAGGTCATA  | TGCAATCATC | AAACCATCAA | ATGATGAGTC | CAAAAGCAAA | 240 |
| 45 | TATGCATTCA | TCAAAATCATC | AAATGAACCA | AAGTAACAAA | AAAGTTTTC  | CAGCTGCTGG | 300 |
|    | TGAAAGTATG | ACATCAAGTA  | TTCTTACTGC | AAGTATTGCC | GCACTACTAT | TAGTATCTCG | 360 |
|    | GTATTCTTAA | GCAATTTAGAC | GACGTTCAAC | AAATAAATAC | ATTAATAATA |            | 420 |
| 50 | GAAAAATCGT | GTGATTATCT  | GaGGGAGCCT | AGGCACATAA | TCAATTGCTC | AGGCTCnCTA | 480 |
|    | ATGGTATATT | GGCAGTAGTG  | GATCAAGTAA | ATTTGGCTGT | GTAACAGCT  | TTTCAATTTC | 540 |

CGATTGCTTT ATACGTGTCA GTGTTAATTC AGATATTTCC TGTGGAATAT ACCACTTATT 660  
 AATCATAATT GGATAAGGTG tTTGTGCGTA CAGTGTTCa ATAATCAGCC AACAAATGTT 720  
 5 ATCACCATCA AACACGTGAC TATGATTTTt GAAGTGGGGC GCTTTGGTAA TAGACAATTT 780  
 TAAATCTGAT TGATATGCAT TGCTATAAAT CGTTTGCTCA ACGAATGTCT TCATGTGCTC 840  
 TTGTTTTTGT GTATTCACCT TAAATGTGTC AATGACATTT AACGGTATAA AGGTAAAGCA 900  
 10 AAATGCATCA GCTTGCTTAG AATGATTGTC CTTTTTTTGA TAATAGCGTT CCATTGCAAT 960  
 GACGCGAGAA GGATGGTTTG CAAACAAATG ATTTGTATAT TCACTTTCTA AATCAACACG 1020  
 ATAATTAATT GATGACATAG ATACGCGAGC TAGCAATATT TGATCAAGTG GATGCTTAAA 1080  
 15 TTGATCCATA CTTGAAGCGT GTTGGGCATT TGTTTGTGGA ATAACAAAGT GTCCCTTCCC 1140  
 TCTTGTACTC TCTACGATGC CATCTTCGGC TAACAATTTt ATAGCTTGGC GCAAAAGTCAT 1200  
 20 ACGACTGGAC ATCAAAAGCG GCACAAAGTT CCTTTTCAGT AGGTAATGC 1249

(2) INFORMATION FOR SEQ ID NO: 390:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1788 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 390:

AATGATGATT CATTCAAAAGA AATTAACGCT CGGTATATGC TTGTTATTAC TCATTATATT 60  
 GATTGTAGGT TATGTCATTA TGACAAAAAC AAATGGTCGA AACGCCCAAA TTAAAGACAC 120  
 35 ATTTAATCAA ACATTAAAAAT TATATCCAAC CAAAAATCTC GATGATTTTT ACGATAAAGA 180  
 AGGCTTTCGA GATCAAGAAT TTA AAAAGGG TGATAAAGGT ACTTGGATAG TTAATTCTGA 240  
 40 AATGGTAATC GAGCCAAAAG GTAAGGATAT GGAACGAGA GGAATGGTGC TCTATATCAA 300  
 TCGCAATACT AGAACCACAA AAGGGTATTA TTTTATAAGT GAAATGACAG ATGACAGTAA 360  
 CGGACAGCCA AAGGATGATG AAAAAAGGTA TCCGGtAAAA ATGGAACATA ATAAAATCAT 420  
 45 ACCAACGAAG CCACTACCGA ATGACAAAGTt AAAAAAGAG ATTGAAAACCT TTAAGTTCTT 480  
 TGTACAATAT GGCAACTTTA AAGATATTAA TGATTATAAA GATGGTGATA TTTCATATAA 540  
 TCCTAATGTA CCAAGTTATT CGGCAAAATA TCAATTGAAT AATGATGATT ATAATGTCCA 600  
 50 ACAGTTaAGa AAAAGATATG ATATTCCAAC CAAACAAGCG CCGAAACTAT TATTGAAAGG 660  
 CGATGGAGAT TTA AAAAGGTT CATCCGTAGT TcTAGAAGTC TTGAATTAC CTTTGTGCAA 720

## EP 0 786 519 A2

|    |   |      |
|----|---|------|
|    | AGSTATGAGT CAAACTGAAT ATCAAATAAA ACCTGGTAAT ATAACAAGTA ACTCTGAAGA   | 840  |
|    | AACAAGTTTCG ATATCTAAAG TGAGCTGTGA AATATAGGTA GCCATTTCAA AAAATTTCAAA | 900  |
| 5  | GGTGAATTTCG ATAATGTAGC TCAAGGAGAT TGGGTTAAAA AGGCGAAGAA TGAACGGAT   | 960  |
|    | GATATTAGTA AGAAATTAAA AAATATTCAA AGAACGGAAG TTTAATGCT TATATGATTC    | 1020 |
|    | TTGGAGCTAA GACAGCATGC GTTCATTTCAT GCCATTATTA ATATAAGCAC CGCAACAAAA  | 1080 |
| 10 | AAGCTTCTAA TGTGATACAG GAACCTCATA TTCCGTATCA TGTTAGAAGC TTTTAATGTC   | 1140 |
|    | TAAAGAACAT CTACATTTTA TCATATTTTC TGACTTATTA AACTTTTATA TAATTAAATA   | 1200 |
|    | TTTCTTAATT TTCCAAAATA GTGATAAATT TGTGAAATAC ATCACAATC CCTTTATTTA    | 1260 |
| 15 | TTTGAAATT CATGTAATAT TAGACTTGTA AGAAGTTAAT AAATAGAGAG AGACGAGAGA    | 1320 |
|    | GTTTATATAA ATACTATATA AACATTGGAG TGATGATTAT GAGAAAAGAG ATTGAAGCGC   | 1380 |
|    | TTATTTCTC AGACGTAAGA GCTATGATAT TTACGTGAAC ACTGGTGTA ATCAAGGATT     | 1440 |
| 20 | AATTGGTGAC ATCAAGATG GTTACCTAAC TATTGATTCT ATGCCTTACA TTGATGCTGA    | 1500 |
|    | GCGTTGTAT CACTTTGCTA TGGAACTGTA ATGGTTAGTC ACTAACTAGT TCTTATTGCC    | 1560 |
| 25 | AATGATTACT ACCCCTAGTC GCGGCAATT GAAGTGTGAT TGATGTAAT TGCCCTCGTT     | 1620 |
|    | GGTGAGCAAT TGAGGCAGA CCCCTTAAAT TAAGTAAACC CTAACCTCCC ACAAACTCGT    | 1680 |
|    | AACGATACTA AAAGCCACGT CCTATATTGG ATGTGGCTTT AGTCAKACTT ATATTATTTT   | 1740 |
| 30 | TAAACGATT ACCTACAAGA TTTACATATA AAATTCATC ATGnCTGC                  | 1788 |

## (2) INFORMATION FOR SEQ ID NO: 391:

|    |                               |
|----|-------------------------------|
|    | (i) SEQUENCE CHARACTERISTICS: |
| 35 | (A) LENGTH: 2407 base pairs   |
|    | (B) TYPE: nucleic acid        |
|    | (C) STRANDEDNESS: double      |
|    | (D) TOPOLOGY: linear          |

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 391:

|    |  |     |
|----|--|-----|
|    | GCAAGTTTAG TTAACAGAT ATTAACAACA GATCAATATA CAGATGCAAA TTCTCTATTA   | 60  |
| 45 | GAGATTAGTT TGCAACAGG TATGTTTATT GCAGGTGGAT TATCAGAAAT ATTATATAAA   | 120 |
|    | ATAGAAAGGAT TCACCTTAAT TATAGCGATG ACTATAATGA TGTTTCTAAT CAGCATTTTT | 180 |
|    | ATGTTATTTA GATTGCATGT AGATAAACCA ACTCATTGAG AGGAAGAATC AACAAATAGC  | 240 |
| 50 | TTATTGCAAG AATATTATT AGGATGGAAA TTTTAAAAG ACAACATGAT GATATTATT     | 300 |
|    | TTTGGAGTTA TTTCAATCAT ACCAATGGTG TTTACAATGA TCTTTAACAT ATCATTACCA  | 360 |

55





GGCTACTATG AAATCGATGA ATGGAAAGAT TGGTTGAAGC ACGTCTTCGA ACAAGACATT 2280  
 GCCAATTTAT CTATTCATCT TCGTACACGT AAAGAAATGA GTAAAGTAGA TGCACATTGG 2340  
 5 GAATTAATCG AAGCTATTAA AAATTTACGT GACGAAATTG CACCAAAATC ATTGTTAACA 2400  
 ATTAACG 2407

(2) INFORMATION FOR SEQ ID NO: 392:

10 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 2424 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 15 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 392:

20 ATGATGhATG GChCGCCAA GAAGTTGAAC CAGTCTATTG GCTAGAATTG CTCAAAAAGC 60  
 GAGAcATGTG GTATTCaTAT GTTaGTAgCt ACgCAAGAC CATCTGICAA TGTAATTACA 120  
 GGTTTAATTA AAGCCAACAT ACCAACAAGA ATTGCATTTA TGGTATCATC aAGGTAGAT 180  
 25 TCGAGAACGA TATTAGACAG TGGTGGAGCA GAACGCTTGT TAGGATATGG CGATATGTGA 240  
 TATCTTGGTA GCGSTATGAA TAAACCGATT AGAGTTCAAG GTACATTTGT TTCTGATGAC 300  
 GAAATTGATG ATGTTGTTGA TTTTATCAAA CAACAAAGAG AACCGGACTA TCTATTTGAA 360  
 30 GAAAAAGAAT TGTGAAAAA AACACAAACA CAATCACAAG ATGaATTATT TGATGaTGTT 420  
 TGTGCATTTA TGGTTAATGA AGGACATATT TCAACATCAT TAATCCAAAG ACATTTCCAA 480  
 ATTTGGCTATA ATAGAGCAGC AAGAATTATC GATCAATTAG AGCAACTCGG TTATGTTTCG 540  
 35 AGTGCTAATG GTTCAAAACC AAGGGATGTT TATGTTACGG AAGCAGATTT AAATAANGAA 600  
 TAATTATGAG TAAGGAGTTT TATATAATGA CACACTATCA TTTTGTCCGA ATTAAGGGTT 660  
 CTGGCATGAG TTCATTAGCA CAAATCATGC ATGATTTAGG ACATGAAGTT CAAGGATCGG 720  
 40 ATATTGAGAA CTACGTATTT ACAGAAGTTG CTCTTAGAAA TAAGGGGATA AAAATATTAC 780  
 CATTTGATGC TAATAACATA AAGAAGATA TGGTAGTtAT ACAAGGTAAT GCATTTCGCGA 840  
 45 GTAGCCaTGA AGAAaAGTA CGTGCACATC AATTGaAATT AGATGTTGTA AGTTATAATG 900  
 ATTTTTTAGG ACAGATTATT GATCAATATA CTTCAGTAGC TGTAACGTGT GCACATGGTA 960  
 AAACCTCTAC AACAGGTTTA TTATCACATG TTATGAATGG TGATAAAAAA ACTTCATTTT 1020  
 50 TAATTGGTGA TGGCACAGGT ATGGGATTGC CTGAAAGTGA TTAITTCGCT TTTGAGGCAT 1080  
 GTGAATATAG ACGTCACTTT TTAAGTTATA AACCTGATTA CGCAATTATG ACAAATATTG 1140

55

EP 0 786 519 A2

|    |   |      |
|----|---|------|
|    | TGGCACATAA TGTAAAAAA GGTATTATTG CTGGGGTGA TGATGAACAT CTACGTAAAA   | 1260 |
|    | TTGAAGCAGA TGTTCCAATT TATTATTATG GATTAAAGA TTCGGATGAC ATTTATGCTC  | 1320 |
| 5  | AAAATATTCA AATTACGGAT AAAGGTACTG CTTTGTATGT GTATGTGGAT GGTGAGTTTT | 1380 |
|    | ATGATCACTT CCTGTCTCCA CAATATGGTG ACCATACAGT TTTAAATGCA TTAGCTGTAA | 1440 |
|    | TTGCGATTAG TTATTTAGAG AAGCTAGATG TTACAAATAT TAAAGAAGCA TTAGAAACGT | 1500 |
| 10 | TTGGTGGTGT TAAACGTCGT TTCATGAAA CTACAATTGC AAATCAAGTT ATTGTAGATG  | 1560 |
|    | ATTATGCACA CCATCCAAGA GAAATTAGTG CTACAATTGA AACAGCACGA AAGAAATATC | 1620 |
|    | CACATAAAGA AGTTGTTGCA GTATTTCAAC CACACACTTT CTCTAGAACA CAGGCATTTT | 1680 |
| 15 | TAAATGAATT TGCAGAAAGT TTAAGTAAAG CAGATCGTGT ATTCTTATGT GAAATTTTTG | 1740 |
|    | GATCAATTAG AGAAAACTACT GGCGCATTA CAGTACAAGA TTTAATTGAT AAAATTGAAG | 1800 |
| 20 | GTGCATCGTT AATTAAAGAA GATTCTATTA ATGTATTAGA ACAATTGAT AATGCTGTTA  | 1860 |
|    | TTTATTATAT GGGTGCAGGT GATATTCAAA AATTACAAAA TGCAATTTTA GATAAATTAG | 1920 |
|    | GCATGAAAAA TGCCTTTTAA TATGTTTATA ATAGAGTAGT ATGGGTTATT ATTATTAATG | 1980 |
| 25 | ACATTATTAC ATGTTAATTA GGAGGCGTTT TTAATGGATT GGATTTTACC AATTGCTGGA | 2040 |
|    | ATTATCGCTG CGATTGCAIT CTTAATTTTA TGTATCGGTA TCGTAGCTGT ATTAATTTCT | 2100 |
|    | GTTAAGAAAA ACTTAGATTA TGTGCAAAA ACACTTGACG GTGTAGAAGG TCAAGTTCAA  | 2160 |
| 30 | GGTATTACTC GTGAACAAC AGATTTACTT CATAAAGTAA ACCGTTTAA C TGAGGATATC | 2220 |
|    | CAAGGTAAAG TAGATCGTTT AAACCTCAGT GTAGATGCTG TTAAGGTAT CGGTGACTCA  | 2280 |
|    | GTACAAACGT TAAACAGCTC TGTAGATCGT GTAACAAAT CAATTACACA TAATATTCT   | 2340 |
| 35 | CAAAATGAAG ATAAATCTC ACAAGTTGTT CAATGGTCAA ATGTTGCAAT GGAAATTGCA  | 2400 |
|    | GACAAATGGC AAAATAGACA CTAC  | 2424 |
| 40 | (2) INFORMATION FOR SEQ ID NO: 393:                               |      |
|    | (i) SEQUENCE CHARACTERISTICS:                                     |      |
|    | (A) LENGTH: 738 base pairs  |      |
|    | (B) TYPE: nucleic acid  |      |
|    | (C) STRANDEDNESS: double  |      |
| 45 | (D) TOPOLOGY: linear  |      |
|    | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 393:                        |      |
| 50 | ChATAATATT ChtCAAnCCT TTAAATAGAG GAATAGACTG CCGACAGAGT CCGAGACTTT | 60   |
|    | GTGGGTAGTT TTTTATGTTT TGATAACGGA AGTTAGAGGC TCTCTGTCAA ATTGGGCAGA | 120  |

55

AGGTAAAGT TITCAACATA ATACTATTAG TTCGGTCATG TATCGGACTG ATGAAAAAGC 240  
 GTTTCACITT TAATGACTCA TTAAGAACGG CCTGAAAATG TTGGCGTAT TAAGTGCAAT 300  
 5 GATAGTTTTG ACATTIAGTT TCTAATTGGT CATTACTGCC GAGCAAAATCT AGTAGAGTAA 360  
 TCATGTAAT CTTTAATGTG CCATTGTGATT CACTAGCGGT GTTAATAACT ACGGAAATTTG 420  
 CATTTCGCAC TGAAATTTTT GAAAAATATC AACGTACGCT ACAAATAAAA TTTTAAACTG 480  
 10 TTATAAATGT GTCTCAAITT CATATGTTCA TCGACGATAT GAAGCGTATT ATGGTAAAAAT 540  
 GAAGAAATAA TAAACTTGTT AATAAATAAA ACATCAGCAT TTGACTAAAG CACTTTATTA 600  
 15 TTGTGTAGAT AATAGTTTTT TAACGAAATA AAAATGGCGA CTGGTTTTAA TAAATCAGCT 660  
 AATGAATCAC TACACCTATA AGTATGAATA TAGTGATTAG AATGCTTTGT ATAGTTGGAT 720  
 TTTCAAAAAT TGATGTTA 738

## (2) INFORMATION FOR SEQ ID NO: 394:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1270 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 25 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 394:

30 AAAAGTTGTA ATTAAGAGTG GGATTTTACT TAAGnAGAA GGAACTATT TATATGACTA 60  
 ATAAAAGAGn AGATGCCGC AATATAGCAA TTATTGCTCA CGTTGACCAT GGTAAAACAA 120  
 CTTTAGTAGA TGAGTGTGTA AAACAATCTG GTATATTGAG AGAAAAATGAA CATGTCGATG 180  
 35 AACGTGCAAT GGACTCTAAC GATATCGAAA GAGAGCGTGG AATTACGATT CTAGCCAAAA 240  
 ATACGGCTGT TGATTATAAA GGTACACGTA TTAATATTTT GGATACACCA GGACATGCAG 300  
 40 ACTTTGGTGG AGAAGTAGAA CGTATTATGA AAATGGTTGA TGGGGTTGTC TTAGTAGTAG 360  
 ATGCGTATGA AGGTACAATG CCTCAACAC GTTTTGTACT TAAAAAAGCG CTAGAACAAA 420  
 ACCTGAAACC TGTGTGTGTT GTTAATAAAA TTGATAAACC ATCAGCACGT CCAGAGGGTG 480  
 45 TTGTAGATGA AGTTTTAGAT TTATTTATTG AATTAGAAGC AAAGnTGAA CAATTAGAAT 540  
 TCCCTGTTGT TTATGCTTCA CGAGTAAATG GTACAGCTAG CTTAGATCCT GAAAAGCAAG 600  
 ATGATAAATT ACAATCATTATATGAAACAA TTATTGATTc ATGTACCAGC TCCAATTGAT 660  
 50 AACAGTGATG AGCCCAATAC AATTTCACAG TAGCATTGTT GGACTACAAT GATTATGTTG 720  
 GACGTATTGG TATTGTGCGT GTATTCAGAG GTAAAAATGCG TGTCGGAGAT AATGTATCAC 780

# EP 0 786 519 A2

GATTAAACG TTTAGAAATT GAAGAAGCAC AAGCTGGAGA TTTAATTGCT GTTTCAGGTA 900  
 TGGAAGACAT TAATGTTGGT GAAaCTGTAA CACCACATGA CCATCAAGAA GCATTGCCAG 960  
 5 TTTCTAGTAT TGATGAGCCT ACTCTTGAAA TGACATTTAA AGTTAAACAAT TCTCCATTTG 1020  
 CTGCGCGTGA AGGTGACTTT GTAACAGCAC GTCAAAATTCA AGAACGTTTA AATCAACAAT 1080  
 TAGAAACAGA TGTATCTTTG AAAGTTTCTA ACACAGATTCC TCCAGATACA TGGGTAGTTG 1140  
 10 CTGGTCGCGG TGAATTGCAT TTATCAATCC TTATTGAAAA TATGCGTCGT GAAGGTTATG 1200  
 AATTACAAGT TTCAAAACCA CAAGTAATTA TTAAGAAAT AGATGGTGTA ATGTGTGAAC 1260  
 CATTGGAACG 1270

## (2) INFORMATION FOR SEQ ID NO: 395:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1365 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 395:

AAGACCAGGA GAACaGTAA AACAAATATAC AGTTGaATC GCTCGTAAAT TAATGGAATT 60  
 TGATATAAAA TGCTCGTGAT TGCTTGTAaa TACGCAACTG CTGThGCTTT AGAATATTTA 120  
 30 CAAAAGACCT TATCAATCCC CAGTGATTGG CGTAATTGac CAGGTGCTAG AaCAGCAATA 180  
 ATGACTACTA GAAATCAAAA TGTATTAGTA CTAGGAACGG AAGGCACAAT TAAATCTGAA 240  
 GCATATCGTA CGCATATTTAA ACGTATCAAT CCACATGTAG AGGTACATGg CGTTGCCTGT 300  
 35 CCAGSTTTTG TGCCACTTGT AGAACAAATG AGATATAGTG ATCCAACAAT TACAAGCATT 360  
 GTCATTATC AAACACTGAA ACGTTGGCGT AATAGTGAGT CTGATACTGT CATTTTAGGA 420  
 40 TGTACCCACT ATCCATTGCT CTATAAACCT ATCTATGATT ATTTTGSTGG TAAAAAGACA 480  
 GTGATTTCGT CTGGATTAGA AACGGCTCGT GAAGTTAGTG CATTGCTAAC ATTTAGTAAT 540  
 GAACATGCAA GTTACTACTGA ACATCCAGAT CATCGATTTT TTGCAACAGG TGATCCTACT 600  
 45 CACATTACTA ACATTATCAA AGAGTGTTTA AATTTATCTG TCAATGTGGA ACGTATATCA 660  
 GTGAATGACT AGGAGGATTT TTAATGAAAG AGATTGTIAT TGCATCGAAT AATCAAGGGA 720  
 AATAAATGA CTTTAAAGTA ATATTTCCAG ATTACCAAGT AATAGGTATT TCAGAACTAA 780  
 50 TACCAGATTT TGATGTGGAA GAAACAGGAT CAACATTGTA AGAAAAATG ATATTAAAAAT 840  
 CAGAAGCTGC TGCAAAAGCA TTGAATAAAA CGGTCATAGC TGATGACAGT GGACTAGAAG 900

EP 0 786 519 A2

CGGATGAAGC AAATATTGAA AAATTATTAA ATAAGCTTGG TAATACAACT GATCGTCGTG 1020  
CGCAATTGCT TTGTGTCATA AGTATGAGTG GCCCTGATAT GGAACAAAA GTATTTAAAG 1080  
5 GTACTGTTTC AGGTGAAATT GCAGATGGAA AATATGGCGA AAATGGTTTC GGATATGATC 1140  
CGATATTTTA TGTACCGAAA TTAGATAAAA CCATGGCTCA ACTTCAAAA GAACAAAAAG 1200  
GGCAAAATTAG TCATAGACGA AATGCGATT AATTACTTCA AGCTTTTCTT GAAGGTGATA 1260  
10 AAAATGTCTA AATGGATTAT TGTGAGTGAT AACCATACTG AATCAGGCGT TTTATATCAA 1320  
ATTTATGAAA TGCACCCAGA TGCAGATGTA TATTTACATT TAGGA 1365

(2) INFORMATION FOR SEQ ID NO: 396:

15 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1383 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
20 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 396:

25 AATTCTCTGGT GCAATAATAA ATAGGATGAA AAAGATTCGG AAAATATGAT AACTCOTAAT 60  
CATAGCAACA TCGGCACCAG TAGCTAATGC AACTAAACT ATCTGATTAA CCCCTCCTGG 120  
TGCTGCACCA AGAAACAAIT CATTAATAGG ATTATTATCA AAGAAATGTA TGATATAAAC 180  
30 CATGATTAGC GCACCAATTA TCAACATAAT ATTTTGAATT GTAATTGCGA TTGCTAGTCT 240  
ACCTTTTAAA TCTGACAATA AATGCGCAAT TTGAACCTCA ATTCTAATCA TATATATTAG 300  
TTGTGCCATG TTCAACAACC AATGATCTAG TGTAAATGTT AAACCTGTAG AAAAATTCCA 360  
35 AACAATTAAT ACAATGAGTG GTGCTAATAA TTGAAATGTT GGAACCTTTA TTTTAGACAT 420  
AATTAGATAA ACTATAAAGA TAGCTATCGC TAAATAACT ATTTCGCCCTA TGTTTAATAC 480  
TTGTGATAAA GGCAAGACTT TTGTTAACTT TCCATTGCGA TGCATGTTAT CATCATGAAA 540  
40 AAAATATGAA ATGAACGGTA CTAACAACAC AACAAATATA ATTCTGTATG TTTGCGTTAA 600  
GCTAACAACT AACAAATTAG CACGTTTGTG TTGTTGAGCC ATGACCAGCA TTTGTGTTAG 660  
TGCTCTGTGT ATAACACTTA AAATAGCTGT TTCTGTATTA ATAGTGCAG TTTTTTAAAA 720  
45 AACAAATGCC ATTACTATTG CAATTAATAA TATCGAAATA GATACAACA TAATCGAAAG 780  
CCAATTGTTT TTAATATCCA TAACGACATT TTTGCTAAAC GTTGATCCGA TTTGCACACC 840  
TAATAGTACA ATACCTAATT CACTAAGTAA GAATGGCCAT TTAATATCAA GTTTGAAAC 900  
50 TTTTACACA ATGATTGATG CGATAATAGG ACCAAACATA AATGSAAGTA ATAGTGGCA 960

55

EP 0 786 519 A2

TATCATTGCC ATGTTTTCa CTCTTTCAA TAAAAATAA AATGACTAAA TTGCTGCTTG 1080  
 AGCTTCACGT TTGTTAAGAT AACAAATACC GCTAGCAGTT tTGACTACAA AGCATATATG 1140  
 5 GaCTTTCACt ATCAAGTCG CGCCCATGCC TTATATACAT TTA AAAAGAG CCTGAACAAA 1200  
 GTTCAGGCTC TCAATTtGTC CGTATATTTA TTTTACAATA CGACTTAAAG CCGTATCAAA 1260  
 TGCTTGAATC GTTTTTCaAT ATCTTCTTC GTGTGTGCCG TAGATAAGAA TGTACCTTCA 1320  
 10 AATTGAGATG GnGGnAAAA CACACCTCTT TgNCATTCTC GGTACATTTC TGCAATAATT 1380  
 TCC 1383

(2) INFORMATION FOR SEQ ID NO: 397:

15 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 415 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 20 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 397:

25 TCCACTAAna TGATTACAAT TGCATTAGTT TGGTGGAGTG CATTTACAAT CTTAACGGGT 60  
 ATGATTAAAG ACCACGGTTT AATTaTTTAG TGAGATTCTT ATTTGGTGTA GGTGAGGCGC 120  
 CAATGTACCC TTCTAATGCT GTGTTTAATT CATTTTGGTT CTCTAAAAAT GAAAAAGGTA 180  
 30 GAGCATCAAG TGCAATTATTA GCAGGATCAT ATTTCGGACC TGTATTAGCA CCAATAGTTA 240  
 CAATTGCTAT TGTTAACGCA TTAACTGGC AAGCAGTATT TTACATTTTt GGTGCAGTAG 300  
 GTATTTTAAT GCKGTATtA TGGCGGATTA TTGCCAAGA CTTACCTGag CrACATAGwa 360  
 35 TGGTTAATGA AGCGGAGAAA CGTTTCATTA TGGAAAAATCG TGATATCGTA GCTAC 415

(2) INFORMATION FOR SEQ ID NO: 398:

40 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1141 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 398:

45 TTTTAGaTaa aTyCAATtTt CyATaCTaaa TgATTnCTCT ATTACGTCAA TtCGCTTTT 60  
 50 aTTTTATCGT AATCTTTCCa CTGCAAGCT AAAGCTTCTC CTATTCTAAG ACCAGAATAA 120  
 AATAACAGTC TAGTTAGCTG ACGAGAAGTA TCATTtGTGA TTTGTtCTAC TTTTTCATCA 180

EP 0 786 519 A2

|    |             |            |             |             |            |            |      |
|----|-------------|------------|-------------|-------------|------------|------------|------|
|    | AATGTGGGGT  | CGATAAAGAG | CTTGTAAATGC | TTTTTGGCGT  | AATTGATAAC | TGCTTTAAAA | 300  |
|    | CCTGCCACCA  | CAGATCGTGC | ATAGTCAACA  | GAAAGACCTG  | CATCGTTTAA | CAAAATAATC | 360  |
| 5  | CTGAAAGCAG  | TACATTGCGT | AGTAGTGATT  | TGCCCAATAG  | GGATATTTC  | GAACTTTTCT | 420  |
|    | TTTATGTGAG  | TATTATATTC | TGTAGTTCGC  | TTTTCATTG   | AGCGTGCAGA | AAGATTTTCA | 480  |
|    | TTTTTTAAAC  | GATCAAAAAA | TATATATTCA  | AAGGGTTGAT  | TGTCGCGAGT | TCCATATCTG | 540  |
| 10 | ACATTTTGTA  | TAAATTCGCT | TTGAGCTAGT  | TGGCATCTCT  | TCTTACGTTT | AAACCCACGC | 600  |
|    | TTCATTTTTC  | GTTTGTATAT | ACCGTATACA  | TCTTATATC   | TAATGGAAAA | ATACCATTTA | 660  |
|    | CCTGTATTAT  | ACTCCTTATA | TACTGGCATT  | TGCTTCTCC   | CTCTCAAAA  | TGGGCAAAAA | 720  |
| 15 | ATAATAAGGG  | TAGCGGGGCT | ACCCGAAATT  | TAGTACTAGG  | TACTAAATGT | GATATAATAA | 780  |
|    | AATAAAAAGT  | AGGTGATGTT | ATGACATTTA  | AAAACAATCA  | TAATTTCAAT | GAATTAGTTT | 840  |
|    | TACGAATGA   | AGACATTAGA | ATTTTAAAAA  | ATGTCTTAGA  | AGATGCAGTC | AGTGTTTATG | 900  |
| 20 | ATGAATATTTC | GGTATGTAAT | GAAGAATCCG  | ATTTTGCTTA  | CTGTTTATTA | AGAGACTTAT | 960  |
|    | ATACATTAGA  | CAGCTTAGCT | ATTTTCGTC   | CAATAATGTT  | GAAATATCG  | TAATGACTCT | 1020 |
| 25 | CGATTTTAAAT | ACCATGCATA | ATAGAGTTTC  | TGTTGTTCAAT | AGCAGCTTTG | ACTGAATGTC | 1080 |
|    | TTAAATGTTT  | TTCTATTAAA | TCGTTGTTTK  | CCAATTCG    | GTAAAAATGT | CyTATATTC  | 1140 |
|    | T           |            |             |             |            |            | 1141 |

30 (2) INFORMATION FOR SEQ ID NO: 399:

(i) SEQUENCE CHARACTERISTICS:

- SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 706 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 399:

|    |  |     |
|----|--|-----|
| 40 | ACTGTGGTAG GTTTTTTATT TTGAAGTATT AATCATAACA GACTAATAAT CATGAGGTAA    | 60  |
|    | CTAATAACAC ATATTAACT TGTATCTTA AACTGGTATA ATAAATTAT GTTGAATGA        | 120 |
| 45 | ATATTGTATG ACAGGGGATT CACTTTTATT AAAAGGTAAA ATTAAATAAA GGTTTATAG     | 180 |
|    | AACGTATTTA AATATATGAG GAGTAAACAA ATGGCTGATA GAACGAATAA AGAAATTAAA    | 240 |
|    | ACAGGACGCT TTGTTGCAAC TGCATCAATC GTATTCTCAA TATTATTGAT TATTCATTAC    | 300 |
| 50 | TTTGTTCCTG TGAATATGC GACTGCCAAA CACTTACTTA ATTTAAGCAA TCAAATACACT    | 360 |
|    | TCAGATCAAG CGATTGATTA CATTCTTAAAC AGCTTTAGAT TCACCTGGTAT TATGATATATT | 420 |



EP 0 786 519 A2

ATGTTTCAG TTTATGTATC AATAGTTTG TTTACGTTGA TTAATTTATC AATCACAATT 540  
 CAAGCAATAA AAGTCGCACA GGGTGGTAC TTAACATTGC CAATTTTAAT TGTATTATA 600  
 5 GGTTCGGTTC CATTAGCGAT TTATATGCTT GTTGTTCCTA TCAACGTAA AAGTACATT 660  
 AATCGCTAGA AAATTGATTT TAACAATAAA AATATGAAAA AAAAnn 706

(2) INFORMATION FOR SEQ ID NO: 400:

10 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1187 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 15 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 400:

20 ACACAATCTG AAGATTACG TTGTGGTGCT GGACATGATC GAAAAATTAG AGCTGAACAA 60  
 ATGAGAGAAA TCAGTGATTT TGTAAAAAG AAAAATATCC CTAAGATGA AACGGTATAT 120  
 ATAGGTGGCG ACCTTAATGT CAATAAAGGC ACTCCAGAGT TCAAAGATAT GCTTAATAAC 180  
 25 TGAATGTAA ATGATGTICT ATATGCAGGT CATAATAGCA CATGGGACCC TCAATCAAAT 240  
 TCAATTGCGA AATATAATTA CCTAATGGT AAACCAGAAC ATTAGACTA TATATTTACA 300  
 GATAAAGATC ATAACAACCC AAAACAATTA GTCAATGAAG TTGTGACTGA AAAACCTAAG 360  
 30 CCATGGGATG TATATGCGTT CCCATATTAY aCGTTTACAA TGATTTTTCA GATCATTACC 420  
 CAATCAAAGC CTATAGTAAA TAGTGCTCAA CTAACATAA ACTTGCTCG TTCTAAAAGG 480  
 35 ACGAAGCGAG TTATATTGTT AAAATTGAA TTGACTTACA TTTTAATAAA ATCATCTTAA 540  
 CAACCTTAAT TTTTCaTTAA TACAaGTCTT TACTCTACAC TCAAAcNAGA TTCATACACT 600  
 GCACGTCATA ATAAATCTAT CTATTCAAAT ATAAATAAAA GTTACCTACT ACATTCTATG 660  
 40 TAGCAGGCAA CTTTTATTAC TTATTTCCTT TCATTATCAT TAAGTACTTT TACAAACTTC 720  
 ACATTATGTG TCTTCCAATC AACTTCATAT AATGCTGATA ATTTTCTTC TTTTTATCT 780  
 ACATGGTTTT CACCAGACCA ATAGCCCCAG AAACCATGGC GATTCCAATC TATTTTAAAC 840  
 45 TCATCCATTG ATCTTTTATA ATGAACAACA AATTGTGATT TACCTTTGTC TTTTTATCA 900  
 TGTGACATAA CAGCTAAAAA TTCTGGATTA AACCCCTCAG ACACAGTTAC AGGCATTTTG 960  
 50 TCTTTAGGTG TGAATATATC TTTGCGCCAT AAATTTCAT TCGTGTATA AGAAAAAATT 1020  
 TCACCTTTAG TTCTATTATC ACTATCATTA GTTAATTGTC TCGTATGGTC ATGCTCCATA 1080  
 TTATTTATCA AATGTGCTTC TACTTTCCAA CCTACACCTT TAIGTGACGT AGATTGATCA 1140

55

## EP 0 786 519 A2

## (2) INFORMATION FOR SEQ ID NO: 401:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 847 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 401:

|  |     |
|--|-----|
| CCAGAAITAT TTTTCAAAA AGGACAAATT AACATGTCG ATAACGTTAT CATAGCAAA     | 60  |
| CCGATGAAAG GGACAAATGCC TAGAGGTAAA ACGGAaGCTG AAGATCAACA GTATTATAAA | 120 |
| ACATTGCAAA CTTCTTCGAA AGATCGTGCA GAAAATGTCA TGATTGTGA TTTACTAAGA   | 180 |
| AACGATATAG GGAGAAATATC ACAGAGTGGC TCAATTAAAG TGTATAAACT ATTTTTTATT | 240 |
| GAGGATATA AAATGTGATT TCAATGACT TCGATGGTAA GTGGAACTTT AAAAAATAAT    | 300 |
| ACAGACTTAA CTCAAATTTT AACATCGTTA TTTCTTGTG GTTCGATTAC AGGTGCACCG   | 360 |
| AAACTGAATA CAATGAAATA TATTAAACAA TTAGAAAGTT CACCTCGTG TATATACTGC   | 420 |
| GGACAATTGG ACTATTACTT CCAACTGAAG ATGATAAAAT GATTTTAAAT ATTCGATTTC  | 480 |
| GCACATTGGA GTATAAATAT GGACAAGCGA TTTATGGAGT CGGAGCAGGT ATTACAATTG  | 540 |
| ATTCTAAGCC AAAAGATGAA GTGAATGAAT TTTACGCAAA AACCAAGATT TTGGAGATGT  | 600 |
| TATAATGCAA TTATTTGAAA CAATGAAAAAT TGATAATGGA CATATCCCTA GACTTACTTA | 660 |
| TCATACTAAT CGCATAAAAT GTTCTTCTGa GCGATTAAAC TTAAATTTTG ATGAACATGC  | 720 |
| ATGGCGAAAT GAATTAAACG ATGTAACAAC AAGTATCAC AGTGGTCAAT ATAGACTTAA   | 780 |
| AATCGTATTA AATGCTGAAA GCAAATTGGA AACGATAGTG TCACCTTTAC CTGAGAAAAA  | 840 |
| TAGTTTT  | 847 |

## (2) INFORMATION FOR SEQ ID NO: 402:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 740 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 402:

|  |     |
|--|-----|
| TGAAGATGAA GCAGAAGCAG AAGACATCT GCTACGAGTA CAATCGAAGA AAGAAGAAT  | 60  |
| GTATCGTCGA TTACTTGCTT CTAACCTAAC AAGCGTTCCT GAAAGTTAT CATCATGAAA | 120 |

GTTAACAGT TAAAGATAA AGTGICTAAA ATTGTGATTG AAATGAATAC ATTTGAAGAT 240  
 GAAGCAAATG ATGTCTCTGT TAATGCTGTT TATGCAGAGA AATTAATTCA ATATGGAAAT 300  
 5 AGATATCGTA AGGACTATAG CAATGTTGAT AAGAGCTTAA ATGAAGCTGA ACGATTATTT 360  
 AAAAAATAATC GCTATAAGCG TCGATTGAA ATTGCAGAGC AAGCTCTTGA AAGTGTGAG 420  
 CCAGGTGTTA CTAACATAT TGAAGAAGAA GTTATTAAGC AATAGAACT AGTATGTAGT 480  
 10 TATACTTAAA TAATATGAGC ACTCTGTCAA ATTGGACTGA TGAGTTTAAAT AATTGAAGTT 540  
 AGCCAACGAT ACGTGTCTA GCTCTTTTT TATATGGATA AATGAAAGGG ACAAATAATA 600  
 TAAATAGCAA TTGTTTAAAG ATAAACGTAA TCAAATGTGT TGTTTTAATT AATATAAGTA 660  
 15 GTGAAAAAAG CATAATCACA CAGCTGTTTA AATAGAGTGA AATAGTCTAA TTCTTATTTA 720  
 ATAAGTAGAA ATAAGATTAT 740

(2) INFORMATION FOR SEQ ID NO: 403:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 630 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 25 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 403:

30 ATGCCAATAA TTAACCATG TAATAAATAT ACGTATAGCG TACGACTACC AATATAAGTA 60  
 TATAATTTTT TCTTTGTTGA CATTAAATTT AGAAACGCAG TCATTGCGAT TAATATAATT 120  
 CCATATAATA TAAGTCGTTT AAAAGGACTG AATATACTCT GTCCTTCATT TTCAAGTGAA 180  
 35 GTATATGGTG AACTTCCCAA TAACCAATCT GCATTGATAG GATGAATCAC GTAAACGATA 240  
 AAAAACAAAA TAAAGGTAAT GATAGATACT GGTATTAGTT TTTTATTTTT AAAAAATGCC 300  
 40 GTATGTTTTT TGGTGAAAAAT GTAACCTAGA TAAATATTG GGAAAAATAC GATTGTCCTT 360  
 GAAATGCTTA AGTAGCTATC GATGTTATCT GAAAAACCTG CTCCAATAGA TATAATAATT 420  
 GAACTGATA GCACCTTATA TGGATTAAAT CTCTAACTA TTACTAAAAA GACATGAAAG 480  
 45 AAAAAATAGCG TGATCAAAAA CCATAACGCA AATACTGGGT TAAAGGATC AAGTTGTAAT 540  
 TCGTCACATT TACCTGTTAA GAAATAATAA ATTGAAAAGA ATGCAAAAAA TATCATATAA 600  
 GGTACTATCA AACGTTTGA AATTITTTCT 630

(2) INFORMATION FOR SEQ ID NO: 404:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 6254 base pairs

55

EP 0 786 519 A2

(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 404:

|    |             |             |             |             |             |             |      |
|----|-------------|-------------|-------------|-------------|-------------|-------------|------|
|    | ATTTAAACGT  | TCAGTTTCTA  | AAAGTGGAA   | AATCAAGAA   | GTACGTAAAC  | GTGAATTTTA  | 60   |
| 10 | CGAAAAACCA  | AGGCTAAAA   | GTA AAAAGAA | ATCAGAGCT   | GCACGTAAAC  | GTAAATTCAA  | 120  |
|    | ATAATTAATA  | CCTCTGTGA   | CTCCCTCAAC  | ACGAATATTA  | ATTATATAAA  | ACAAACATCA  | 180  |
|    | CAAGTTAGTG  | TCTGACACTA  | ATATGTGATG  | TTTTTTTGT   | GTCAATTTT   | AATTA AAAAA | 240  |
| 15 | AGTTATATAG  | TTTATAAATA  | ATCAAATTTGA | TATTTCTATAG | GTCTTTATATA | CTATAAAGTA  | 300  |
|    | TATTTCAATTT | CATGTATAAT  | TAATGTGAGG  | CGGAGGTGAA  | ATTGTGAGAT  | ATAATAATTT  | 360  |
|    | TTTACAAATG  | ACCAATCTCT  | TTGTAATCAAC | GCGTGAGTAA  | AGTTGGGTGT  | AACCAAGTTAG | 420  |
| 20 | CAATATAAAT  | GTCTCAACCTA | TTTTTACGTT  | AATATTAAC   | TGTTTGACAT  | TCTTAGGATT  | 480  |
|    | TGTATATCAA  | CTTTACTCTA  | AAAAAATCAA  | TGCAGCTGGT  | ATTATCGCCA  | CATTATCATT  | 540  |
| 25 | ACTTATTTTA  | TTTTTGGGAT  | TTCTAATCCA  | AGGAAATGTC  | AATATGCATT  | CTATCTTAAT  | 600  |
|    | ATTCTCAATT  | GGCGTTATAT  | TAGTTGTAAT  | TGAATTTATT  | GTAGTTGGTG  | CAGTAATTTG  | 660  |
|    | TATTATTTGC  | ATGATACTGA  | TAACATAAAG  | CATAACAACG  | CTCGGTGATA  | ATTTGCTATT  | 720  |
| 30 | TATGCTTGGC  | AATGTTATCG  | TTGCCTTGAT  | TTTAACGATT  | GTAGAATGGG  | TGATATTAGT  | 780  |
|    | GAAGATTTTC  | AACAGAAAGA  | TTCCGTTTTT  | GGATAAAGTT  | ATCTTAAAG   | ATTCAACTAA  | 840  |
|    | TTCTGAGTCA  | GGTTACAATT  | CTCATGTAA   | CCGCTCGCAC  | CTCGTAGGAA  | AGACTGCTCA  | 900  |
| 35 | AACAGTTACA  | GATCTTCGAC  | CTCGAGGGAT  | TATTTTTTGT  | GA AATGAAC  | GTATTTGATC  | 960  |
|    | TGTTTCAGAT  | GGCAACTTTA  | TTTTTGCGCA  | TAAAAAGTTA  | AAATACCTTG  | AAGTTGTAAGG | 1020 |
|    | AACAAGAGTA  | TTTGTGAGGG  | AATGAGATTA  | ATTAAAAAGG  | AGGATACCAT  | TTTTAGTTTA  | 1080 |
| 40 | AGTTTTATCG  | TAATAGCAGT  | TATTATAGTA  | GTTCGATTAC  | TTATTTTATT  | CTCATTTGTA  | 1140 |
|    | CCCATTGGTT  | TATGGATTTC  | AGCGTTAGCA  | GCTGGCGTTC  | ATGTTGGTAT  | AGGTACATTG  | 1200 |
| 45 | GTTGGTATGC  | GTTTACGTCG  | TGTATCTCCA  | AGAAAAGTTA  | TAGCGCCATT  | AATTAAGCG   | 1260 |
|    | CACAAGCAG   | GACTAGCATT  | AACAACAAC   | CAATTAGAAT  | GCCATTATCT  | AGCAGGAGGA  | 1320 |
|    | AATGTTGACA  | GAGTTGTGGA  | CGCTAATATT  | GCTGCACAAC  | GtgcTGACAT  | TGATCTTCCT  | 1380 |
| 50 | TTCGAACGTG  | CTGCTGCAAT  | TGaCCTTGCA  | GGACGTGACG  | TATTAGAAGC  | GGTTCAAATG  | 1440 |
|    | TCTGTTAATC  | CTAAAGTCAT  | TGAAACACCA  | TTTATCGCAG  | GTGTAGCAAT  | GaACGGTATT  | 1500 |
|    | GAAGTGAAAG  | CCAAAGCTCG  | TATCACAGTT  | AGAGCTAATA  | TTGCTCGACT  | TGTTGGTGGT  | 1560 |

55

|    |   |      |
|----|---|------|
|    | GCAACGTATA TATGTTAATG CCATGAAAA TAATGATTTA GTATTGGTA TAGGTCTGCG   | 3480 |
|    | TGGTACAGGT AAGACATCTT TAGCTGTAGT TTATGCAGCA AAGCAACTCC GTAAAGGTGC | 3540 |
| 5  | TGTTAAACGT ATTGTATTAA CAAGACCTGC TGTGGAAGCA GGAGAGTCAC TTGGAATTTT | 3600 |
|    | ACCAGGAGAT TTGAAAGAAA AGGTAGATCC ATATTTAAGA CCTTATATG ATGGTCTATA  | 3660 |
| 10 | TACTGTTCTT GGGCGTGAAC AAACAGAGCG ATTTATTGAA AGAGGCATTA TCGAAATAGC | 3720 |
|    | GCCACTTGCA TATATGCGCG GACGAACATT AGAAGATGCA TTTGTAATTC TTGATGAGGC | 3780 |
|    | GCAGAATACG ACACATGCGC AAATGAAAAT GTTTTTAACA AGACTAGGTT TTGGCTCAAA | 3840 |
| 15 | AATGGTAGTT ACTGGTGACC AAATCAAAT CGATTACCT AAAGGTGTTA AAAGTGGACT   | 3900 |
|    | TAAGGAAGCG GTCAGTAGGT TACACAACGT TAAAGGTATA AGTATATTGA AATTAGATCA | 3960 |
|    | GAGCGATGTA GTCAGACATC CATTGGTAAG TAAGATCATT GAACATTATG AAGGAGAGAA | 4020 |
| 20 | TTAAATGTTT ACGATAGATT TTAGCGATCA CACAGGCTTA GTTAAAGATG CTTGGTATAA | 4080 |
|    | ACAAATGAA GATTTATTAG AATTTGCTAA AAAAGAAGAG CATATAGAAG ACGATGCTGA  | 4140 |
|    | GCTTTCTGTT ACATTGTAG ATAACAAGA AATACAAGAA ATTAATCGAA CATATAGAGA   | 4200 |
| 25 | TAAFGWTAAR GTTmCaGATG tAaTCyCaTT tGCTTTAGrA GrAGATGAGC CmGaGATkGA | 4260 |
|    | TtTAGTGGT CTTGATATAC CACGTGTTTT AGGGGATATA ATTATCTGtA CGGATGTAGC  | 4320 |
| 30 | GCAAGFACAA GCAACAATTT ACGGACATTC TTTTGAACGA GAATTAGGAT TTTTAGCATT | 4380 |
|    | ACATGGATTT TTGCATCTAT TAGGTTATGA TCATATGACT GAAGCGGATG AAAAGGAAAT | 4440 |
|    | GTTTGGTCGA CAAGATACAA TATTAAACGC ATATGGATTA ACACGAGACT AATTATGAAA | 4500 |
| 35 | AGGTTTAAAT ATGCACCTGA TGGGCTGAAA ATCTTAATTC AAAAAGACTA TAAATTTCTT | 4560 |
|    | TTACATGTGT TTGCAATGAT TGTGCTATT GTCTTTGGTC TCGTACTAAA TATTAATCGG  | 4620 |
|    | ATTGAGTGGa TATTTATACT CATTGCTATT GCATTAGTTC TCACTGTTGA AGCTTTAAAC | 4680 |
| 40 | ACTGCTATTG AATATGTTGT CGATTTAGTG ACCGTTGAAT ATCATGATTT AGCTAAATAC | 4740 |
|    | GCTAARGATA TTGCGGCTTT TAGTGACTTT ATAGTTTCAA TATTAGCATT TATTATAGGT | 4800 |
|    | TTAATAGTAT TTTTACCACA TTTTATAGCG TTATTTTAGG GAGGCATATA TGAGTTATCA | 4860 |
| 45 | ACCTCATTAT TTTCAAGAAG TTAGAAAAGC ACAACAAGAA TCATATTCGC CATACAGTCA | 4920 |
|    | ATTTAAAGTA GGGGCTTATT TAAAAmCGAA AGAcgGTAGA ACTTTTATG GTACCAATGT  | 4980 |
| 50 | AGAAAAAGCT TCTTATCCAT TATCGATATG TGCTGAACGA GCTAGTTTGG TATCGCAAT  | 5040 |
|    | TTCTCAAGGA TACAGACCAG GTGATTTTGA ATCAAtAACT GTAACCGTAG ATGCAGATAA | 5100 |
|    | ACCGTCATCA CCTGTGGTGT CATGTCGTCA AGTTTTGAAG GAATTATGTG ATGATGATAT | 5160 |

|    |             |             |             |            |            |            |     |
|----|-------------|-------------|-------------|------------|------------|------------|-----|
| 45 | GTTGTTCTTAA | ATGTTTCTTG  | TATGAACCAAG | GTCAATGTAA | ACTGATATTG | CTGTATTTGT | 60  |
|    | TGCAGCAAT   | CATATTGGTC  | TGGTGTGCA   | ATAACACGAG | CTTGAGTTGG | AGTCGCACT  | 120 |
|    | CTGATGTCTG  | CAGCAAAAATC | ACTTAAATGA  | AAGTCCTGTT | CATGACCAAC | TGCTGATATA | 180 |
| 50 | ATCCGGTGCT  | TACAATTATA  | TATTGACCGG  | ACGACAGCTT | CTTCGTTGAA | ATTCCATAAA | 240 |
|    | TCTTCTATGG  | ATCCACCGCC  | TGCAGCTACA  | AATAGTTT   | CTACACCTAA | ACTACTCTGA | 300 |

EP 0 786 519 A2

|    |            |            |            |            |            |            |      |
|----|------------|------------|------------|------------|------------|------------|------|
|    | TGTATTGTT  | CAGCTAATGG | AAAACGACTA | TTTATCGTTG | AATGGATATC | TCGAATTGCG | 420  |
|    | GCACCTGTAC | TCGCTGTAA  | AACTGCAATT | TTTTAGGAA  | ACTTAGGTAT | TGATTTCTTA | 480  |
| 5  | TTGCGTTTAT | CAAAACAACC | TTCTTCAGTT | AATTTTTTCT | TTAATGCTTC | TAATTTTGA  | 540  |
|    | TATAAGTTCC | CTATACCATC | TAATTGCATT | TTATTTACAT | AAATTGATA  | GTTTCCACGA | 600  |
|    | CGTTCAAAA  | CAGAAACAG  | TGCTTCTAAT | AAGACTTCAT | CTCCTTCTTT | AGGTTGGAAG | 660  |
| 10 | TTTAATTTAG | AAGCACTACC | TTTGAACATC | ATGGCACTTA | TAACGCTTTC | TTTATCTTTC | 720  |
|    | ACATTAAAGT | ATAAATGACC | ACTTGaATGc | TTTTTGAAAT | TTGAAAGCTC | ACCTTTAATC | 780  |
| 15 | AATACAGATT | GGAGATGTGG | ATCTTGATCA | AATTTATATT | TAATATATTT | CGTTAAAGCT | 840  |
|    | GAAACACTTA | AATAATCTGA | CATATAACAT | CACCTCAATT | TATTTTTTTA | TATTACTCAA | 900  |
|    | TACACCATTT | ATAAATTTAT | AATGATCATC | ATCACTGAAT | TGTTTGTGTA | ATTCAACTGC | 960  |
| 20 | TTCACTTATT | ACGACTTTAG | CAGGTGTATC | ACTGTGTAAT | ATTTTATATG | TTGCCATTCT | 1020 |
|    | TAAATAATA  | CGATCCGTTT | TTAATAAACG | TGCAATAGTC | CAATCTTTTA | AATAAGGACT | 1080 |
|    | AATTGTCTCG | TCTAATACAG | GTTCTGTATC | TTTAACGCCA | GAACTAGCC  | AATGAATAAA | 1140 |
| 25 | TTGGAAGTCT | AAATCTGGAT | TATCGTCTTT | AATAAAGCTT | ATCGCTTCAT | TTATCGTTAA | 1200 |
|    | ATCACTGTCC | TTCACTTCTA | ATTGAAATAA | AGTTTGAAAA | GCTTGCATCT | GGGATTCTTT | 1260 |
|    | ACGACTCATT | TTTAACTCCT | TCAAACGTTT | GTATTTTCTT | TTATTTAATT | ACTGAATTAG | 1320 |
| 30 | GTATGACATT | ACTTTTCAAT | AACGATTGTG | GTAATGIGAA | TATTAATTGG | CTTAGGTTCT | 1380 |
|    | ATCGCTGTCA | TATTAGAAAT | TGAATTAAAA | ATTGACGTTT | GAATTTTGTT | TGCAGTTTTT | 1440 |
| 35 | GAAATATTAA | CACCATGTTT | TAATGCACAA | TATACATCTA | TATATATGCC | ATCTTCTTTA | 1500 |
|    | CTCTCGATTT | TTAAATCACG | GCTTAAATTT | TTACGACTAA | CTTTTTCTAA | ATTGTTTCT  | 1560 |
|    | TTTAATTCAG | CAAAATGGCC | AGTGATGCCT | TCGACTTCCG | AAGTAGCTAT | ACTTGCAATA | 1620 |
| 40 | ACAGATAGCA | CTTCTGGCGC | TATTTCTACT | TTACCTAATT | TTGAATTTGA | ATAATCAGTT | 1680 |
|    | ACTTTGACCA | TGGATTGACC | TCTATTAAAC | CTTCATCATT | CATAATGCTA | TTTTGCTCTA | 1740 |
|    | AAAAGTTTGT | ATTAAATTTA | CCGCTTCTAA | ATATATCGTT | ATTCAATAAT | TTAATATGGA | 1800 |
| 45 | ATGGAATAGT | TGTATCAATA | CCAAGAACCA | CAAAATTCAT | TAGTGCACGA | ATGCCAGCCA | 1860 |
|    | TAATCGCTTC | ATCTCGTGTG | GGTTCAATGA | TGATTAATTT | CGTACCATC  | GAATCATAAT | 1920 |
| 50 | ATGGCGGTAT | CGTATAATTA | GTATAACATG | CTGACTCTAT | TCGAACACCA | TATCCACCTG | 1980 |
|    | GTGCAAGATA | TTGCTCATTT | TTACCTGGTG | ATGGCATAAA | GTTCCTGTAA | GGATTTTCAG | 2040 |
|    | CATTAAATCT | AAATTCATTT | CGGTGTCTCT | TTAATTTAAT | ATCTTCTTGT | TTATACGGTA | 2100 |
| 55 |            |            |            |            |            |            |      |



EP 0 786 519 A2

|    |            |             |             |             |             |            |      |
|----|------------|-------------|-------------|-------------|-------------|------------|------|
|    | CAGTTACAGG | ATGTTCTACT  | TGAATACGTG  | TATTCATTTC  | CATAAAATAA  | AATTTATTAT | 2220 |
|    | CATTTAAATC | ATATATAAAC  | TCAATTGTTC  | CCGCATTTTC  | ATAATTTIACA | GCTTTCGCTG | 2280 |
| 5  | CACGAACATG | GGCATTTCCT  | ATTTACAGCAG | GTGTTTCATC  | ATCTAAAAAT  | GGGGAAGGTT | 2340 |
|    | CTTCTCCAC  | TAATTTCTCG  | ATACGTCTTT  | GAATTGTACA  | ATCACGTGTC  | CCTAAATGAA | 2400 |
| 10 | TTACATTACC | ATAGCTGTCC  | CCAAACAATT  | GGATTTCAAAT | ATGCGCGAAG  | TTTTCGATGA | 2460 |
|    | ATTTCTCCAT | ATAAAGTCCA  | CCATTACCAA  | ATGCAGTTTG  | AGCTTCTTGT  | TCTGTCATT  | 2520 |
|    | GGAAGCCAGT | TTCAAGTTCT  | TTTTCATCAC  | GAGCAACACG  | GATACCTTTT  | CGGCCACCCG | 2580 |
| 15 | CAGCAGTAGC | TTTAATGATG  | ACCGGATAGC  | CAATTTTTTT  | GGCGATTTTC  | TTAGCTTCTG | 2640 |
|    | AGACGCTCTT | CATTAAACCG  | TCACTAACAG  | GAACAACATG  | AACATTGGCT  | TTGATCAATT | 2700 |
|    | CTGCCCTTTT | AACATCTTTG  | ATACCCATTT  | TTTGGATAGA  | TGTATACTT   | GGTCCAATGA | 2760 |
| 20 | ACTTCAATTG | GcATGctTCG  | CATAAATCTG  | CAAAATCAGC  | ATTTTCAGCT  | AAAAAGCCAT | 2820 |
|    | AACCCGGATG | AACGCCATCA  | CAACCTGTAG  | AAGTTGCAAT  | AGATAAGATG  | TTCCGAATAT | 2880 |
|    | TTAAATATGA | ATCTTTTAGAC | AAAGTGGGAC  | CTACGCAATA  | TGCTTCATCA  | GCAATTTGAG | 2940 |
| 25 | TATGTAGCGC | ATCTTTATCC  | CCTTCAGAAT  | AGATTGCAAC  | AGTTTGGATG  | CCTAAATCAC | 3000 |
|    | GACAAGCGCG | AATAATCCTA  | ACTGCGATTT  | CACCGCGGTT  | TGCAATTAAG  | ACCTTTTTCA | 3060 |
|    | TTATTTACCC | TTAAATAACG  | GTGGCCATA   | CTCTACCATT  | TGTCGGTCTT  | CTACTAAGAT | 3120 |
| 30 | TTCAACAATT | TCACCTGAAA  | TTTCTGCTTG  | AATTTCAATTA | AATAGTTTCA  | TGCGCTCTAA | 3180 |
|    | AATACACACT | GTGGTTTCAT  | TTGAAACAGT  | GTCCCCAACT  | TGCACATATG  | CTTCTCGCTC | 3240 |
| 35 | TGGAGATGGC | GATTTGTAA   | ATGTACCTAC  | CATAGGTGCA  | TTAATTGTTT  | TGTGATTATC | 3300 |
|    | TGAAGTTGGC | TTTGAGAGCTT | CAGTTTTATT  | GCTATCAGTT  | GATTGTGCTT  | GAGGCATAGG | 3360 |
|    | CATTGCGGCA | GCTTCAACCTG | GCAATTTGTA  | GATTTGTGGC  | GTGATAATCT  | CAGTTTCTTT | 3420 |
| 40 | TTCTTTCTTA | AGCGTCACCTT | TGCGTTTAGT  | ATCTTCAATA  | TTGATTTCGG  | TTAAAGTTGA | 3480 |
|    | TTTATCCAGA | ATTTCAATTA  | ATTCCTTTGAT | TTCTTTAAAG  | TTCAATTATTA | CTGACTCCTT | 3540 |
|    | CAGTTTGGTT | TCATCTACCC  | GTCTATTTTA  | CTTGAAGACA  | CTCTTCAATT  | CAAGCATGTT | 3600 |
| 45 | CATATTGCTG | GGCAGACATTT | AGTCTATATC  | CCTAGATATA  | ATAAAACCAC  | ATTTTAAATT | 3660 |
|    | AAAAACACTT | GTGATTTTAT  | TACTTAACAT  | TGACTCATCT  | TAACTCTTGA  |            | 3720 |

(2) INFORMATION FOR SEQ ID NO: 406:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1705 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 406:

|    |   |      |
|----|---|------|
| 5  | GCTGATGTTT GTTGCCCTTn TCCACCAGAC AATTCAGAGG GATATTTATC ACTAATATCC | 60   |
|    | AATATATTTA ATGCTTCTGC TACTTTTTCA TAACGATTTA ACATATGTTT TTTATCTAAC | 120  |
|    | TTCTGTACCG TTAGTGTAA CATTATGTTT TCTTTAACAG TCAATGTATG CAGTAAATTA  | 180  |
| 10 | TACTCTTGAA AATAAAACC AATATCATGC TTGGTATAT CAGATWATTC CTGTGTTGAA   | 240  |
|    | AGCTTTCTTA ATTTTTTTC TTTTAATGTA ATAGAACCTT GTGAAATATA ATCAATTGAA  | 300  |
|    | CTTAAACAT TTAATAATGT CGTTTCCCA GATCCAGAGG GACCATAAT AGCAATAAAC    | 360  |
| 15 | TCGCCCTCTT CAATAGACAT AITGATATCT CGCAACACTT CTGTGCCAT TTTTTAGTT   | 420  |
|    | CCATATATTT TTGTTAATTG TTTTACTTCT AAAATTGCCA CTTTAACACT CCTATAATT  | 480  |
|    | ATCTTAACCT CATTTCCCTT AGGCTTTGGC ACTTGATATC TCAATTTAAC ATATGACTAA | 540  |
| 20 | CATCTATCTT ATTATAACGT TGAAGCTGCA TTGATGTATC AATCTAAGT AACAAAACGC  | 600  |
|    | ATGTTTAAAA TGACAAATTT GTCACCTCCG ACATGCGTTC AACAAATTCA TTTTGTAATG | 660  |
| 25 | GGAAATCAA TCTGACGTT GTCCCTTAC CAACAGTCGA CGTGACTTGC AGGTGAATAC    | 720  |
|    | CTAATGTATC CTTTACACTA TTTACTAAAT ATAGACCCAT ACCTGAAGAC GTCGTTTCAT | 780  |
|    | TTCTGTTAGC CGTTGACGTA AATCCTCGTT CAAATATTGC CGGCATATCT TTTTACTAA  | 840  |
| 30 | TACCTCTGCC ATAGTCTTTA ATATATAACG AAACATGTTG ATCATTTAAT TCTGTCCCAA | 900  |
|    | TTTCAATATT AAAATTCTCA CTATATTICA ATGCGTTTGA CAAAATTTGT CTAATAATCA | 960  |
|    | TACGACACCA TTTTATATCT GTATAAACAT AATCATCCAC TTTAAAGTCA ACATCAAAAC | 1020 |
| 35 | CAATACCTTT AACCTGACTA ATATGTCTTG TTAATTGTAT TTCATCAATG ACCATGCGTT | 1080 |
|    | TAAGTGACAC GTAATCAAAA TACATATCTT TAGCTTGAGA TTCTAATCTA GTAATATACA | 1140 |
| 40 | GCTGTGTATC TAGCATCGAG TTTATACGAG ACCATTCTA TAGTAATGCT TGTWTTCTTT  | 1200 |
|    | CTGTGATTTT TTCTGTATCA ATTAATAATT TCATAGCTGT CACAGGTTT TTTATGCTGT  | 1260 |
|    | GCACAAATTC TGTAAATGTT TGTTCATGCA TGTTCAAATG CAACGTGTGC TCAACAACT  | 1320 |
| 45 | TTTCTTTGTG CGCTGAGATT TGACGATATA AATAATCAAC TGTATGACGT TGAATGGCG  | 1380 |
|    | TTCCGCTAA ATCTTTAATG TTAATTTCTT CTATTTCTTT ATCTTTGTCA AAATGCTTAT  | 1440 |
| 50 | ATAATTTTAC TTCTTTAAAA TATGTCAATA AAAGAAAAAT CATTGTTAAA CTTAAATTCA | 1500 |
|    | AAGAAACAAT ATAAAAATAA CTGTCTATTG GAAATCATA ATCGATTAGA CTATGCCATA  | 1560 |
|    | ACATAAGGAA GTTTAAAAAC AATATCCAAA ATATCCAGTT CATGCGAGAT TTCAAAAAAT | 1620 |

55

GCACTATCCA TACTAATTTC AGATA

1705

(2) INFORMATION FOR SEQ ID NO: 407:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1722 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 407:

|    |  |      |
|----|--|------|
| 15 | TCAATATATC TTAAAAATTCA ATGATTAAAT CnATTATCAC TAGACATnAA ATACATAAAT | 60   |
|    | CCTATTCGCC ATTTCATT nTTAATTTCAT AAATGAATCA ATAACCACCT AATAACAAAT   | 120  |
|    | CATATTATAC ACCTTTGTTC TCTATTTTTC TAAGGTTTAA AAAATATTTT TAGGTAAACC  | 180  |
| 20 | TAAAAATAGA TGTAATAAAA ACGCCTCCTC AGATAATTAT ATATCTATGA AGACGTTTAA  | 240  |
|    | ATACATTATA GATGGTCTGG TTCTGGGTGA ACGTATACTG AGGAAALACC TTTTGTGTGC  | 300  |
|    | AAATGATGTT CGACATTGTC ACAAAATTGA TGCCTTCTA CTAAGGAAAG GTTAGCATCT   | 360  |
| 25 | ACAACAATTG TGACATCAAT AAACACACTA CTTCCATGGT AACGCCCTTT AATACTTTTA  | 420  |
|    | ACTTCTTGTA CTTCATCAAC TTCTAAAATA TCATTGCGAT ACGCTTCTAA TTCAGTTTCA  | 480  |
| 30 | TTGAAACCAT CACTCAACAT AAAAATTGCT TCTTTAAAAA TACCAAAACC AGTATAAACG  | 540  |
|    | ATTAGTAAGC CTAGTAATGT TGCTAAAATA ATATCGACAA TTGGGAAACC GATTTCGGTA  | 600  |
|    | AAAATTAATC CTATCGCTGT TCCAATGCTG ACTAAACTAT CCGATAAATT ATCTTTGGCA  | 660  |
| 35 | GCCGAATTTA AAGAACTACT TTTCGTTCTT TTCGCTAGTC TTTGATTGAC TGCAAAATACA | 720  |
|    | ATCAACATTA CAAGACCACT GATTAAGCTG ACGATAATTG TTATTGCGTT AGGTACAACG  | 780  |
|    | TCATCTTCTT TGAACAAACG AGGTGCATTT TGAATAACTA CTGGATACC TACAAACATA   | 840  |
| 40 | ATGACAAATG ACACCAATAA TGAAGAAATA TTTTCAGACT TCAAAATGCC ATAAGGATGA  | 900  |
|    | TTTCGATCGG CAGGTTTAAT TGAATTTTTC AATCCAATAA TAACAGCTAA AGAAACGATA  | 960  |
|    | ATATCTGTCA TATTGTTTAA TGCATCGGCT CTTACAGCTG CAGAGTTAAA GACAAAACCC  | 1020 |
| 45 | GTGACATACT TAACAATAGA TAAGATTATA TATACAATTA AACTCAATAA AGCACGCGT   | 1080 |
|    | TGCGCCAATT TAAGATTTTC ATTATGAGAC ATGCGTTGAA CCACCTTGAA TTAGTATAGT  | 1140 |
| 50 | AACAATATTA TGAATGATTC ATTTTAATTT TACAACGTTT TTAATTTTTA TAAATTTTTA  | 1200 |
|    | TAAATTTAA CTAATTTTAT CATTGCAAC CCCTAAAAAT AATTTTTCAG CTTTCTGCGA    | 1260 |
|    | ATTTTATGAG CTAGAAAGGC GCCCACTCT CCCTGTTTGT TAACTTTTCG CTCGAAAGTT   | 1320 |

EP 0 786 519 A2

CGAAWTTIAT GAGCTAGAAA GGCTTATGCA GTTGACGTTT TACGTCCAAC TGCCTTCCTC 1440  
 CGTCTTCTTC AAATTTATTT GThAGAAAGG CACCCAAC TCCTCGTTTG TTAACCTTGG 1500  
 5 CCTCGAAAGT TTCTATGTTA GAACCCATG CATGAGTTGC GAATATCTA ATGTCGTGAA 1560  
 CTAATTATAT AGAAGAAAAA GTGCATCAAT GACAAATTAA ATGAGATTTC TACTCTACCA 1620  
 AACTCTCTTC GAAAGACAAAT TTCTCTCTCT ATTTATTAGC AACTATTGCA TTCTCCATA 1680  
 10 TAGTACTTCC TTACTTAAAA TACGCTGAAT GTCTGAATTA AA 1722

(2) INFORMATION FOR SEQ ID NO: 408:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 5521 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 408:

GCGAGACCCC CTGAGGGAGC AGTGCCCGT CGAAGACCCG AGGCTGAGAC GGCACCCCTAG 60  
 25 GAAAGCGACc ATTCAATAC GAATTgTGAT AAATAGAGAA CAGCAGTAAG ATATTTTCTA 120  
 ATTGAAAAAT ATCTTACTGC TGTTTTTTTA GGGATTTATG TCCAGCCTG TTTTTGTGA 180  
 30 TTTTAAATAA TTGAATATG GAAAATGTAT TATTCTCTCA TTGTATAGA TTGATATTAA 240  
 TAAGTTAATG TAATCCTTGA GCTCACGATT AATAAAATTC TATAACCTTA ATTATTTCT 300  
 CGATACAAAG GGTTATTAAc TTTAATATAA GTATAATGAT GTGCCTCATC TTCAAGACGC 360  
 35 ATTGTGTGTA CACACTTATA ATCTATAAAT GCGCGGAACA TGSTATCTTT AATTTCaTTT 420  
 ATGCGATTCT CATTTACTTG ATTAGATTG TGtGTCGAAA GTACAAGTTG ATCaAAAAATG 480  
 TTATCTAGTA CATCACGAAC GATATACCAC ATATGTCTTT CTAAGTTTGA ATCATTTGAT 540  
 40 GCTTTAGAGA TTGTAAGAAT TAATTCGCCT AAATGGTTT GAACGGTAGA ATAAAAGGCT 600  
 TTGTAAACA CAGAGTTTT TGAATCAGTA AGAATCTCTG ATTTTTCATG GAAATGAGAT 660  
 GTACTGTATA CCATTTTCATT TAGTTGTGCT TTATCAATTC TTAACCTTC AAAGTCCTA 720  
 45 ATATACATCG TGTCCAATAG ACCATCTTTC CGAAATGTTG CAATAGCATT TTGCAAGTGA 780  
 GCCTCTAATG CAATGCCATA TTTAGTAACT AGTGGGATTA CGAGACCAAG CAATGCTTTA 840  
 CTATAAGTTT CAACCCACGA TTTCGCCGAT GATTCAAAAT CAGACAATGA TGCAGCTGAT 900  
 50 TGATAACGTT TAATCAATGT CACGATAGGT GATTCAATTG TAAATGGGTA GGTGCAACT 960  
 AAGCTTGAAG GAATCAATGG TGTGACTTCT TGtGGAATCA TTTGTATAT ATTTTCTA 1020

55



|    |            |             |            |             |            |             |      |
|----|------------|-------------|------------|-------------|------------|-------------|------|
|    | CAGTTTTATT | ATGCTTACCA  | TTACATTTTA | AAGTAACATA  | AATACCTGAA | GaTGATCAAA  | 2940 |
|    | GaTACATGCC | GTTAAAGATT  | ATTATAGATT | ACTTCAAATT  | ACATATGGAA | GGTCGACAAA  | 3000 |
| 5  | TATTTATAAC | ATCATATTGG  | ATTATGGCGA | CAGGTTTTTC  | ATATACGACA | CTTTTACCAG  | 3060 |
|    | TTTTGACAAA | CAAAGTATTT  | CCGGGGAAAT | CTGAAATATT  | TGSTATCGCT | ATGACGATGT  | 3120 |
|    | GTGCCATTGG | TGGTATTATT  | GCAACGCTAG | TTTTACCTAA  | AGTACTTAAA | TATATTGGTA  | 3180 |
| 10 | TGGTAAATAT | GTATTATTTA  | AGTTCAATTT | TATTTGGCAT  | TGCTTTGGTA | GGTGTGGTAT  | 3240 |
|    | TTCCAAATAT | TGTCATCATG  | TTCAATTGTA | TTACATTGAT  | TGGGTTATTT | AGTCAATGGG  | 3300 |
| 15 | CAGTACGAC  | AAATCCGGTT  | TATTTTCAAA | ATTATGGTAA  | AGGATTATGA | CGTGTAAGG   | 3360 |
|    | TACTGAGTAT | TaTTTATGAT  | GgATAGAGGT | ATGaTTCCAT  | kgGGGAAGTc | ATTAATGAGT  | 3420 |
|    | ATATGTGCAG | ATGTGTTTGG  | CATTGTGTGA | ACTTTTTCAA  | TAATGGGAAT | AAGTACTATA  | 3480 |
| 20 | TGCATTACAA | TGGTATTCTA  | TTTTATAAAT | AGAAAGTTGA  | AGTTAAAGTT | GGAGGAAAGT  | 3540 |
|    | AATCATGTGA | TATCTTGAAT  | GGGCAAAGGC | AGATAGAAAT  | ATTCAATATC | GTGTAATTAA  | 3600 |
|    | CGCCATTATT | AAAGAACGTA  | TTTACCCCCG | GCAAAACATT  | ATTTCGCAAA | AAGGATCTTT  | 3660 |
| 25 | AATTGAAATT | CAGTATCATa  | TGCATGTGTT | GACTATTGAA  | GTGTGTAGAA | AAAGTGCAAT  | 3720 |
|    | AGAACGCTAT | GAGTTTACAG  | GTGATATTAC | TTATTTAAAT  | AAAGGTGAAA | CGTCATTAAAT | 3780 |
|    | TATAACTTTA | GAAAGTTTAT  | TAGACGTGTT | GAATCATGAC  | TTTGATATCC | TTTATTTCAGa | 3840 |
| 30 | CGCACTACCG | GAAAGAGTTAA | TACAGATCGC | AGATAGTTTA  | GTGAAACATC | ATAAGCAAAAT | 3900 |
|    | GTCTCACAGA | CAAAAGTTTAA | TAAgTcAAAG | TTTTAAATT   | TCAGGGTTAC | CACAAGATAT  | 3960 |
| 35 | TAACTTTTTT | TcAkGGTtAC  | AACATGTAAA | AGATAGTGAT  | AAGACAGATG | ATTTAACTTA  | 4020 |
|    | TTCTGAGAGT | TGGTACCAG   | AGGGGCATCC | AACACACCTC  | TTAACCAAAA | CGAAATTGCC  | 4080 |
|    | CTTAACTATG | GAAAGAAGTAC | GAGCATATGC | ACCTGAGTTT  | GAAGAAAGAA | TCCCTTTGCA  | 4140 |
| 40 | AATTATGATG | ATTGAAAAAG  | ACCATGTTGT | GTGCACAGCT  | ATGGATGGTA | ATGATCAATT  | 4200 |
|    | TATTATTGAT | GAATTAATTC  | CGGAATACTA | CAATCAGATT  | CGTGTGTTTT | TAAAGAGTTT  | 4260 |
|    | AGGTTTGAAA | AGTGAAGACT  | ATAGAGCGAT | TTTAGTACAT  | CCTTGGCAAT | ATGATCATAc  | 4320 |
| 45 | GATAGGGAAA | TATTTTGAAG  | CATGcNTTGC | TAAGAAAAATA | TTAATTCCAA | CGCGGTTTAC  | 4380 |
|    | AATACTTcCA | AAAGCAcATT  | aTcAtTtTAG | ACGATGTCTT  | TAATTGATAA | ACcATACCAT  | 4440 |
|    | GTTAAGTTCG | CgCTCGATGC  | aAGTCAACAA | AGTGCcCGTGA | GAAACAGTCT | AACTGTGACT  | 4500 |
| 50 | ACTGTAGATG | GACCAAAAGT  | CAATTATGCT | TTCACAAAACA | TGTTGAATCa | ATATCCAGGA  | 4560 |
|    | TTTGAAGTGT | CTGACGAACC  | GTTCCGGTAA | TGTCAAAGT   | TAGTAAGAAG | TGGGCAAGCT  | 4620 |

AGTGCAAGTC TAGTTAATAA AAATCCAATA GATCAAAAAG TTATCGTGGA TAGTACTTA 4740  
 GAGTGGTTAA ATCAAGGAAT TACTAAAGAA AGTATTACGA CATTATTTGA ACGATACGCT 4800  
 5 CAAGCATTAA TCCCGCCTTT AATIGCTTTT ATTCAAAATT ATGGAATTGC TTTAGAAGCA 4860  
 CACATGCAAA ATACAGTAGT GAACTTGGGG CCACATTTTG AcaTTCAATT TTTAGTGAGA 4920  
 GATTTAGGTG GTTCTAGAAT TGATTTAGAA ACATTACAAC ATCGTGATC AGATATTAAA 4980  
 10 ATTACAAATG ATAGTTTAAAT AGCTGATTCT ATAGATGCAG TGATTGCAAA ATTCCAACAT 5040  
 GCTGTATTC AAAATCAAAT GGCAGAATTA ATCCATCAIT TTAATCAGTA TGATTGTGTT 5100  
 15 GAAGAAACCG AATTATTAA CATAGTACAG CAAGTAGTAG CGCATGCCAT TAACCAACA 5160  
 CTACCACATG CAAATGAGTT AAAAGATATT TTGTTTGGAC CAACAATTAC TGTCAAAGCG 5220  
 TTGTTAATA TGAGAATGGA AAATAAAGTA AAGCAATATT TAAATATTGA GTTAGATAAT 5280  
 20 CCGATAAAAA AAGAGGTGTA GTACTACATG GCACACGTTA ACATAAATAT ATCGAAGATT 5340  
 AAATATAACG CCAAAGTACT TCAAACAGTT TTTCAAAGTA AAAATATGCA ATTCACACCA 5400  
 GTAATTAAAT GCATAGCTGG TGACCGTACA ATTGTAGAAA GCTTAAAGC GTTAGGTATC 5460  
 25 AATCATGTTG CAGAATCCAG ATTGATAAC ATAATTAGTA TTGCAGATAC AGGATTAAAC 5520  
 A 5521

(2) INFORMATION FOR SEQ ID NO: 409:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1261 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 409:

AGGTGCTCTr GCaGmAGmGm TATTIGACGA TCTTGCTTTC CCAAAACAGC ATGATGATT 60  
 TAACATACTG TCTGATTATA TTGAGACACA TGGTGATTTC aCATTGCCAA TGTCGTGATT 120  
 TGATGATTTA TATGAAGAAT ATACGGAATG GCTAAAAATT TAATATAAAT TTTAATAATA 180  
 45 ATAGTTAGAA CCAGGGTGAT GCAATTCGTT ATCTCGGTTT TAATTTAAAA TAACTAAGT 240  
 TGTGACTAAA AATTAATCAA TTATAGTGAA ATATGGTGCG CTATCTTGCA TAAATTGATA 300  
 TGATTAACCTA CACAGAATT AAAAGTACAT AATACATAAT AAGGAAGTGA TACAATGGAT 360  
 50 GATAAGCAAC ACACATCTTC ATCCGATGAT GAACGCGCTG AAATTGCAAC AAGCAATCAA 420  
 GACCAAGAAA CTAATTCATC GAAACGCGTT CACTTAAAC GTTGGCAATT CATATCAATA 480

CAAAAATAA GTGGTTTAAA CAAACTGAT CAAGCAAAT TAAATAAAAT TGAATATGTG 600  
 TATAAAATCT TAAATAGTGA TTATTACAAA AAACAGGACT CTGACAAGTT AAGTAAAGCT 660  
 5 GCAATTGATG GCATGGTCAA AGAATTAAAA GATCCTTATT CTGAATATTT AACAAAAAGAA 720  
 CAAACGAAAT CCTTTAATGA AGGTGTTTCA GGTGATTTTG TAGGTATTGG TGCAGAAATG 780  
 CAAAAGAAAA ATGATCAAAAT TATGGTTACT AGTCCTATGA AGGGATCTCC AGCAGAACGT 840  
 10 GCTGGCATTG GTCTTAAAGA TGTCATTACT AAAGTAAATG GAAAATCAAT TAAAGGTAAA 900  
 GCATTAGATG AAGTTGTCAA AGATGTTGCT GGTAAAGAAA ACACCTGAAGT CACTTTAACT 960  
 GTTCAACGAG GTAGTGAAGA AAAAGACGTT AAGATTAAAC GTGAAAAAAT TCATGTTAAA 1020  
 15 AGTGTGAGT ATAAGAAAAA AGGTAAAGTT GGAGTTATTA CTATTAAATA ATTCCaGAAT 1080  
 GATACATCAG GTGAATTGAA AGATGCAGTT CTAAAGCTC ACAAGATGG TTTGAAAAAG 1140  
 20 ATTGTTTAG ATTTAAGAAA TAATCCAGGT GGACTACTAG ATGAAGCTGT TAAATGGCA 1200  
 AATATTTTAA TCGATAAAGG AAAAATGTT GTTAAACTAG AAAAAAGTAA AGATACTGAA 1260  
 G 1261

25 (2) INFORMATION FOR SEQ ID NO: 410:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 2488 base pairs  
 (B) TYPE: nucleic acid  
 30 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 410:

35 AAATATATTG AAnAGaNAAT TACTAAGATT AAATChTCTT AAAATATCCC TGAaATAAGC 60  
 TCCTAAAGAT TAAAGGAAAG AGGTtATAAG TTATGCCAAA ATTAATTTTA TGtGTCATG 120  
 40 GACAAAGCGA GTGGAATGCT AAAAaCTTAT TTACTGGATG GGAAGATGTT AATTtATCTG 180  
 AACAAAGTAT TAATGAAGCG ACTAGAGCAG GTGAAAAAGT AAGAGAAAAAT AACATTGCCA 240  
 TCgATGTAGC TTTTACATCG TTATTAAACAC GTGCTTTAGA TACAACGCAT TATATTTTAA 300  
 45 CTGAATCTAA ACAACAATGG ATTcCTGTAT ATAAAAAGCTG GCGTTTAAAT GAACGCCACT 360  
 ATGtGTGGAT GCAAGGCTTA AATAAGATG ATGCTAGAAA AGAATTtGGA GAAGAACAAG 420  
 TACATATTtG GCGTCGTTCT TATGATGTGA AACCACCTGC TGAAACCGAA GAACAACGTG 480  
 50 AAGCTTACTT AGCTGATCGT CGATATAATC ATTTAGATAA ACGTATGATG CTTTATtCTG 540  
 AAAGTCTGAA AGATACTTTA GTTCGAGTGA TACCATTtTG GACAGATCAT ATTTCACAAT 600

55



|    |  |      |
|----|--|------|
|    | TTAAATATCT TGAAGATGTG TCAGATGAAG ATATCATTAA TTATGAAATT AAAACAGGTG  | 720  |
|    | CACCGCTTGT TTATGAATTA ACGGATGATT TAGAAGTTAT AGATAAATAC TACTTATAAA  | 780  |
| 5  | AaAAGAGCTG CATGTACACA AGGAGTGAGT GTATATGcAG CtcTTAAALa TGTGAAGTAA  | 840  |
|    | TGTAAAGGAA TAGTTAAGTA TAGAGTTTAT ATTAACGAGC TAGGGATACT CGAAATATATA | 900  |
|    | GTTAGACATA CAATATAGTC AAATTAaAAC AATTATTTCG CTCTTTTATG TTGCTTAATA  | 960  |
| 10 | ATCTTTAAAG CACGCTTTCT TGTTTTAATG TTAGGGCTAT TTAAAITACG ACGAGCAGTC  | 1020 |
|    | TGTAAATCTA ATTTTCATCTC TATCCCTCCT TGTAAATATA TTATGACCGA TAACTACTCA | 1080 |
| 15 | TATGTAATA GTAATGATTA CGTTTTAAAG AAATGTGAAT AAAGTCGTGC TAATTTTTTG   | 1140 |
|    | GAAATGGGT ATAATTACCG GATATCTAAA AATGTGTGTC GTTTTTTAGA TGGTAGGGG    | 1200 |
|    | GAAGCTTAA ATGTGGAAGA AACAAAAATT AACGATGATT ATTACTATGC TGATGGGTGG   | 1260 |
| 20 | ATTTTTTGA TTATTAAATG AAACACTATT AGTGACGGCT TTACCAAGTA TTATGAAAGA   | 1320 |
|    | TTTTGAAATT TCATATACAC AAGTTCAATG GCTGACAACA GCTTTTTTAT TGACTAATGG  | 1380 |
|    | GATTGTATT CTTTGTCCG CGCTTGTAT ACAACGTTAT ACAACAAGAC AAGTGTTTTT     | 1440 |
| 25 | AGTGGGTATT TCTATCTTTT TCTTAGGTAC ATTACTCGGC GGCTTGAGTC CGCACTTTGC  | 1500 |
|    | AACATTATTA GTTGCTAGAA TTATTCAGGC GTTAGGCGCA GGTATTATGA TGCCATTGAT  | 1560 |
|    | GATGACAAGC ATTTTGGATG TTTTCCAACC ACATGAACGC GGTAAATATA TGGGGATATT  | 1620 |
| 30 | TGSTTTGGTA ATTGSTTTAG CACCAGCTAT TGGACCTACT CTTTCAGGTT ACCTTGTTGA  | 1680 |
|    | ATATTTTAAc TGGAGATGc TTTTCCATGT TGTGCTCCA ATTGCAGCTG TGACATTTTT    | 1740 |
|    | AATTGGaTTT AAAAcGATAA AAAATGTTGG AACTACAATT AAAGTACCTA TTGATTTTAT  | 1800 |
| 35 | TTCTGTcATT TTTTCTGTAC TAGGTTTCGG cGGGTTATTG tATGGAAcGA GTTCaATTTC  | 1860 |
|    | AGAAAAAGGT TTTGATAATC CTAcGeATTa GTATCTATGA TTGGAGGCGT TGTTTTAGTC  | 1920 |
| 40 | GCATTATTTG TAwTAcGTCA ATATCGGCTA TCAACACCAT TATTaAATTT TGCTGTATTT  | 1980 |
|    | AAAAATAAAC AATTTACAGT TGGTATCATT ATTATGGGTG TcCaAATGGT ATOGATGATT  | 2040 |
|    | GGTTCGGAAA CGATTTTACC TATCTTTGTG CAAAATTTAT TGCATCGTTC AGCTTTAGAT  | 2100 |
| 45 | TCTGGATTAA CTTTATTACC AGGAGCAATT GTTATGGCAT TTATGTCGAT GACTTCGGGT  | 2160 |
|    | GCTTTATATG AAAAGTTTGG TCCTAGAAAT CTTGCTTTAG TAGGTATGGC GATTGTGTTT  | 2220 |
|    | ATTACTACGG CTTATTTTGT TGTAATGGAT GAACAAACAT CAACAATCAT GTTGCGAACA  | 2280 |
| 50 | GTTTATGCGA TTCGAATGGT TGGTATCGCG TTAGGATTAA TTCCAGTAAT GACCATACG   | 2340 |
|    | ATGAATCAAT TAAAGCCAGA AATGAATGCA CATGGTTCAT CTATGACAAA CACAGTACAA  | 2400 |

EP 0 786 519 A2

AACTTTTCTC CAACTATGTC AGACTATA

2488

(2) INFORMATION FOR SEQ ID NO: 411:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1105 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 411:

|             |            |             |             |             |             |      |
|-------------|------------|-------------|-------------|-------------|-------------|------|
| TTTACAAATT  | TCAGATATCT | CTAAACACACA | TACCGATCCA  | ATGGCCAAAA  | TTATTCTGTCA | 60   |
| AAAATTGAAA  | AAACTAGGAA | TTCTGTAAGG  | GATTCAGATT  | GTATTTTCAG  | ATGAAAGTCC  | 120  |
| AATTGTGATA  | AGAGAAGATG | TAAAGATAT   | AGTTGAGAT   | AAAAATGCTA  | TCAAATCGAAA | 180  |
| AGGACAAATG  | CCACCTTCTT | CAAAATGCTT  | TGTGCGCAAGT | GTTGTGTGAT  | TAAATTTGTGC | 240  |
| AAGTTATGTG  | GTGAATGACG | TATTAAAGA   | TATTCAGATT  | CGTCGCATTA  | AAGACAAAGS  | 300  |
| GCAATAATTC  | ATTITGAAAG | GGATAATTTT  | CAACGTAAAG  | CAAGGTGAAC  | CACACCATAA  | 360  |
| AACTAAAACT  | GACTAGTTTG | CAAGCATAGT  | AGAAATATGCC | TCGTGTACTA  | GTCAAGTTTGT | 420  |
| ATTGTGATTAT | AGCTAAAATA | TCAAAGGTAA  | TCCGATTAAAT | CTTATTTTTA  | CGGCCTTTGT  | 480  |
| TTAAGTAAGT  | TGTTATATAT | TTCTTTGAGT  | TGTTGTTTAC  | TTTTAGATGT  | CGTTTTTGGT  | 540  |
| TCGTAATAAA  | TTTTGTTTTT | TAGTTTATCA  | GGTAAATATT  | GCTGTGAAAC  | ATAGCCATTA  | 600  |
| ACATATTGGT  | GTGGATATTT | GTAACCAATA  | GATCGGCCTA  | GATCTTTIAG  | ACCTTGATAA  | 660  |
| TGTCATCTTT  | TTAAATGTTT | TGGAATTTGG  | CCCACATGAC  | CGTTTCTAAT  | ATCGGACAAT  | 720  |
| GCACATATCAA | TGGCAGTCAT | TGCTGAATTT  | GACTTAGGTG  | ATAAGCATAG  | TTCGATTACT  | 780  |
| GCTTGGCCTA  | GTGGAATTCT | AGCTTCTGTT  | AAACCTAGAC  | GTTCTGCTGA  | TTCAATAGCA  | 840  |
| GCAAGTGTCT  | TCTGACACCG | ATTAGGTGAG  | GCTAAGCCTA  | TATCCTCATTA | GCTAATFACA  | 900  |
| CGAATCTGTC  | GAACTATTGT | AGGTAATCT   | CCAGCTTCAA  | TTAATCGTGT  | TAAATAATGT  | 960  |
| AAAGCGGCAT  | TGACGTGCT  | ACACGGATA   | GATTTTGTGA  | AAGCGCTCAT  | AACATCGTAA  | 1020 |
| TGCATGTCAC  | TATCCTTGTC | ACTTACAAT   | GCACCTTTTT  | GTAACACGTC  | TTTAGCATCT  | 1080 |
| TGCAATGTAA  | CATGTCGATA | ACCGT       |             |             |             | 1105 |

(2) INFORMATION FOR SEQ ID NO: 412:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 579 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double

EP 0 786 519 A2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 412:

5 TAACGTCGAA CTGAGCTGT TACGTTATGA CTCATAATTA TTTAGCATA GTCATTAAAA 60  
TAAACTCTGT TTCTGTCTGT TGGATAAGCA AATTCAAGCA ATTGACTGTA ACTATCATTA 120  
ATGACTTCTT GATCAACATG ACTATCAAAA TATACAGCAT AATAATAAGT ACCATCAACC 180  
10 ATATATAACA AATCTTCAAA CTCTGTAGTT ATTGGATTGC TATGATATGC ATAAATTAATA 240  
ACATCTTCTA AATGTTTAAA TTTCACAATG ATTGTTCTTG TATTTTACG TGCTGAAGAC 300  
TTTTGACGTT TAGAACCTTG AGCTTCTTTT TCTTTGTTC GTTGCTCGAA TAATCTTCT 360  
15 AATTGATCTT CACCTTCTAA TGTTTGAGCT AACAAATCTT GAACCTGTTC ATCAAAATkGA 420  
TCAGTTGCAT CATCATCAGA CATATTCTAT ATATCTTCAT TTTTAGATTT AGAAATTTGTG 480  
ACTTCGACAC CTTTTCAAA GGCATGTACT TGAATCCATA ATGGACCTnC TACAACAAA 540  
20 TCTTCTACTT CGTAAATTC ATCCATCATT GAnCAAAAG 579

(2) INFORMATION FOR SEQ ID NO: 413:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1342 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 413:

GTGGTGAAC TTCTTGCTTT GTAATTTTAT AAAGTGGATC AATATCACTC TGACTTACAT 60  
35 CAGGTTGAAC TGTCATCTTT TTAGTTACTT TGTTTTCAG CGTAATATTT GCTAATTTcTA 120  
GGCTTTACAG AATAATGAAT GTCATCAGCT AmCCCTTTmm CtTGATACTy CTCTTGATAT 180  
TTTCCAGCTT TAGCATTGTA TAAATCAATC ACTACTCTTA AATCTTCTGG ATTTTCAATT 240  
40 TTTATTATCT TTGATTGTG TCCTGAAATA GTCACATTAA CTGTTTCAGG CGCTTTTGTT 300  
AAATGCAAAAT CTTTAGTGTT ATAAAGAATT TCAACGGGTA CATCTTGAAT CGTTTIACTA 360  
GACTTTTGAC CAAGATTACC AGTGTTAAAG ATATTCCAA AACATTGTT AACAGATAAA 420  
45 AAGAAAAACA ATGCCAAAAG AAAGGCAATA AATCTCAAGC CCCATTACT TTCTAGCATA 480  
TTATTTCACA CCTTTCTTTT GAAAGCGTGT GCCAAACCAA TGTTCCAGCA GCAACTCTTC 540  
50 AAAAATTCG TTTGAAATGT CTGTCGTAA TTTTCCATCA AATGTTACCG AAATATCACC 600  
AGTTTCTTCA GATACAATAA CGGTAAATGC ATCAGATACT TCTGAAATAC CAACGCAGC 660



TTTAGATACA ATTACGAGTA TTTCAACACC AATGGGTGAA GGGCCAATTG GAATTGTTCTG 840  
 ATGTCTCGGA CCGCAAGCCG TTGAAATTGC TGACAAATTA TATAAAGGAA AACATCTTTT 900  
 5 AAATGATGTT CCATCAGATA CGATTAAATTA CGGTCAATATT ATTGATCCAG TCTAAAGAAG 960  
 TGGTTGAAGA AGTTATGGTG TCTGTGTAA GAGCGCCAA AACATTACA CGCGAAGATA 1020  
 10 TTAGAGAT TAATTGTCAT GGTGGATTK TAACGATTAA TAGAGTGCTG GGA 1073

## (2) INFORMATION FOR SEQ ID NO: 415:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 3176 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 415:

CTTTACCAAT GCCAAATCCG AAGTAAAGTA TAGCAATAAA GATTACTAAT ACAATTCTGT 60  
 AAATGGCAAA TGGAAATTAGT TTGATTTTGT TAATTAGATG CAAGAATGTT TTGATTGCAA 120  
 25 TTAGTCCAAC AGTAAATGCA GCTAAAAAGC CTAAATATA AAAAGGTATA TCAGCAATCT 180  
 GAATATCTTG ATAATGTTTT AATAAGATA AACCACTAGC TGCTAACATA ATTGGAACAG 240  
 CCATAATAAA TGTAAAGTCC GATGCTGCTT TATGATTTAA TTTCAATTAAT ACCCCAGTTG 300  
 30 AAATTGTTGA GCGTGAACGG CTGAAACCAG GCCACATAGC TACTGCTTGA GAAATACCAA 360  
 TTACAAATGC TTGGAATAA CTGATTGAT CTACTGTTTG TGGGTTTTTA ACTTTAGCTG 420  
 AGTATTATC AGCAATAATC ATATAGATAG CACCTACGAA TAAGCCAATC ATAACAGTTG 480  
 35 GCACACTAAA TAAATGTTCT TCGATGAAAT CATCAAATAG TAAGCCTAAA ATACCTGCTG 540  
 GCACCATACC CACTAATACA TGTAATAAAT TTAAACGTCT TGGCTTTGAA CGTCTTTGTT 600  
 40 GATCGTTATC TCCTTCAACA TGTTTGTTT TACCAATATG TAAAATCTCT AAGAAGCGTT 660  
 CCGGGAACAC CCATGCTGCT GCAAGACGG ATCCTAATTG GATGACGATT TTAATGTAA 720  
 ATGCTGACTG AGAACCTAAA AATTCAGATG ATTTTAACCA CATATCATCA ACTAGGATCA 780  
 45 TATGTCCAGT AGAGGAACA GGTGCAAAAT CTGTTAATCC TTCGACGCC CCTAAGATAA 840  
 TACCTTTTAT TAATTCAATG ATAAACATAA TGTACCCACT TTCATTACT AATTTAATTT 900  
 ATTTAAATAT CAAATTAACC ATATCATGAT AGCATATTCA TTTAAAGACA TGCTAGTTAT 960  
 50 AGTTATAATA CTAGACTAAA GATGTATATA TTCATTTTCT TTTACATGTA AAACATAAAT 1020  
 ATTTTATTGA GCTATTAAAT TTGATTTTAA GGAAAACCTT TTATAATAGG TTAGGGTAGT 1080

EP 0 786 519 A2

|    |  |      |
|----|--|------|
|    | TTCTTGGTCA GTACTGGTCT CGGCATAATC GTTATAACGC AAAATATTTT AATAGCAGAT  | 1200 |
|    | TTTTTAGCTA AAATTATAAG ACATCAATTT CAAGGTTTAT GGATTGTATT ATTTATTTTA  | 1260 |
| 5  | TTAGGTGTTT TACTTTTAAAG AGCAACTGTG CAATTTCTAA ATCAATGGTT AGGTGATACA | 1320 |
|    | TTAGCATTTA AAGTTAAGCA TATGCTTAGA CAGCGGGTTA TTTATAAAAA TAATGGTCAT  | 1380 |
|    | CCAATCGGTG AACAAATGAC TATACTCACA GAAAACATTG ATGGCTAGC ACCTTTTTTT   | 1440 |
| 10 | AAGAGTTATT TGCTTCAAGT GTTCAATCA ATGATGGTTC CGCTCATCAT AATCATTCOA   | 1500 |
|    | ATGTTTTTCA TCCATTTCAA TACCGCATTA ATTAGTTTAA TAACGTGACC ATTTATTCCT  | 1560 |
| 15 | TTGTTTTATA TTAATTTTCG TTTGAAAACG CGAGATGAGT CAAAAGATCA AATGACTTAT  | 1620 |
|    | TTGAATCAAT TTAGTCAACG GTTTTTAAAT ATTGCTAAAG GTTTAGTGAC GTTAAAGCTA  | 1680 |
|    | TTTAATCGTA CAGAGCAAC AGAGAAGCAt ATTTaCGACG ATAGTACTCa GTTTAGAAGT   | 1740 |
| 20 | TTAAACATGC GCATTTTaCG CAGTGCTTTT TTATCGGAT TAATGCTCGA ATTTATAAGT   | 1800 |
|    | ATGTTAGGTA TTGAAITGGT TGCATTGGAA GCAACGCTAA GCTTAGTAGT ATTTTCATAAT | 1860 |
|    | ATTGATTTTA AAATCGCGC AATTGCGATT ATTTTAGCGC CTGAATTTTA TAATGCAATT   | 1920 |
| 25 | AAGGACTTAG GSCAAGCGTT CCATACTGGA AAACAAAGTG AAGGTGCCAG TGACGTGTG   | 1980 |
|    | TTTGAGTTTT TAGAACCAACC GAACTATAAT AATGAATTC TATTAAAGTA TGAGGAAAC   | 2040 |
|    | CAAAAGCCAT TTAITTCAGT AACAGACATA TCATTTCGAT ATGATGATTC TGATAGATTG  | 2100 |
| 30 | GTATTAAATG ATTTAAATTT GGAATATTT AAAGGTGATC AAATTGCATC TGTAGGTCCA   | 2160 |
|    | AGCGGGGCGA GTAATCCAC TTTGACACAT CTTATTGCAG GTGTTTATCA GCCAACATA    | 2220 |
|    | GGTACTATAA GTACAAACCA GCGTGATTIA AATATAGGAA TACTTAGTCA ACAGCCATAT  | 2280 |
| 35 | ATTTTCAGTG CTTCTATAAA AGAGAATATT ACGATGTTTA AAGATATAGA AAATAATACT  | 2340 |
|    | ATTGAAGAG TGCTAGACGA AGTAGGTTTA TTAGACAAAG TGCAATCTTT CACAAAGGC    | 2400 |
| 40 | ATTAACACAA TAATAGGTGA AGGAGGCGAA ATGTTATCTG GTGACAGAT GAGACGCATA   | 2460 |
|    | GAACTTTGCC GTCTTTTAGT TATGAAGCCA GATCTCGTTA TATTTGATGA GCCTGCAACT  | 2520 |
|    | GGTTTAGATA TTCAACAGA ACACATGATT CAGAACGTTT TGTTTCAACA TTTTAAAGAT   | 2580 |
| 45 | ACAAGCATGA TTGTCATTGC ACATAGAGAT AATACAATTC GCCATTTACA ACGACGCTTG  | 2640 |
|    | TATATAGAAA ATGGAAGACT GATTGCTGAT GATCGCAATA TTTCAAGTAA TATAACAGAA  | 2700 |
|    | AATGGTGATG ACTTATGAAA ACACGACTAA AATTTCAAGT AGATAAGGAT TTATTGTTAG  | 2760 |
| 50 | CTATAGTTGT TGTTGTTTGT GGAAGTTTAG TTGCGGTCGC CATGTTTTC TTAAGTGGTT   | 2820 |
|    | ATATGGTGAC ACAAAGTGCA CTTGGTGCGC CACTATACGC TCTGATGATT TTAGTCGTTA  | 2880 |
| 55 |  |      |

## EP 0 786 519 A2

ATAAAGCTAC ATTTACAATG CTACGTGATA TTCGGGTACA GTTTTTCGGT AAATTAGTAA 3000  
 ATGTCATTC C TAATGTTTAC CGTAAACTGA GTTCTAGTGA TTTAATTICA CGIATGATTA 3060  
 5 GTCGTGTGA GGCATTACAA AATATAKATT TACGTGTTA TTATCCACCA GTCGTCATCG 3120  
 GTTTGACAGC GCTAGTTACA GTCATAGTTT TGGCGTTCAT TTCAATCGGC CATGGG 3176

(2) INFORMATION FOR SEQ ID NO: 416:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 2109 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 416:

TCCTTTATTC AATTTC AAT TGAATTTTT TATTATTAC GCATAGCTCT TAAATTAAC 60  
 GTTACGATTG CAATTAAGAT AATTGAACCA ATTAATGCTG GCAAGATGTA AATACCTCCT 120  
 AATTACAGGAC CCCATTGTCC GAATAGTTTG CCACCTACCC ATGATCCAAAT AATACCTGCG 180  
 25 ATAATATTGC CTAAATATACC ACCTGGGATA TCCTTACCCA TAATAGCACC AGCAGCCCAT 240  
 CCAATTAAGC CACGACAAT TAACATTCCA ATAAATCCAA ACATAATTTT CAGTCTCCTT 300  
 TTTCTATTTA TTTTGGTTA TTCTAAGTAG TACCCCTTAT TTACAATTCT AAAACAAATT 360  
 30 CAAATTAATT TTATCCAAAT ATTTTAAAA GTAGTAATTG AATATCAATT TTATTC AATG 420  
 TAGCTATCGT TATTTAAAGT CTCTGTACCG ATAATATCAT ATACATTTAC ATTATTTTTT 480  
 CTGCCGAATT CATAGCTTGA TTAATTTTATG TTATAGGACT AgAATATACA CATATTATTA 540  
 35 GAGCATCTTT GAATTTTAAA TCAAGAAGCG AGGTTAATGA ACAATGAATA TGCATATTTT 600  
 ATATAACTTA CGAACTAAAC ATAATTTAGA AATTGACGAA TTAGCACAGC AATTAAATGA 660  
 40 GAAATATGGT ACTAAATATG AAGCACATCA AATTGGGAA TGGGAGAATC ATCACCATGA 720  
 ACCTAAATTT AAAGATGCCA TGCATTAGC TGACTCTTT GATGCACCAT ATGAAATGTT 780  
 TTTAGAAAGT AAGGTTAAAG AATATCAGAA ACATTTAGAA GAAGTCGATA TTGCGCATGA 840  
 45 TAAATAGATG CAAATAAACC CTCACAACAC GTTTGGCATA TATCCTTTCA AATCTATACT 900  
 GGATATATTA CATTACGTTG TGAGGGTATT TTATTAATTA ATATGAATTA AGACATTTTA 960  
 CAAGCGTTAA TGCAACGAAT CTTTTTAGTG ATCTGTGCTA CTCTTTAATA CTTTACGTT 1020  
 50 CTTAGCATCA ACAGTAACCTT CTGTGTTTT ATTACCTTTT TTCAATCGA TATGTGTAAC 1080  
 AAGTTTGCCA TCATCTTTTT CAAGTGACCA TTCTTTAATA TCACCATCAA ATTCTTTTTG 1140

55

## EP 0 786 519 A2

|    |  |      |
|----|--|------|
|    | ATTCATTGTA TCTCTTTTT CAGTCTTTTT GTTAATCACT TTTTATTTT TATCAGCAAC    | 1260 |
|    | AAGTACTTCT GACTCTTCAC CAGATTTTIG TTGGTCACT TTATAAGCCC ATTCACCATT   | 1320 |
| 5  | AGAATTTTCA AATGAAATTC CTTTCAACTT TTGGCCTTG TAAGTTTCTT CAGCTTTTTT   | 1380 |
|    | CACAGCATCT TCTGGGCTTG TTTTAACATC TTTTAAGCA ATAACATCTT TTGTTGTATT   | 1440 |
|    | AGTGTCTTGA TTAGTATTG ACTCTGTGTA TTTTGTTC TCTTTTGGAG TATCATTGGC     | 1500 |
| 10 | ACATGCAGTA AGCACCAACG CTGACATTGA TAACACTGCT AATGATTTTA ATTTCATAAT  | 1560 |
|    | ATCACTCTCT CTTCTATTT TTGAAACTCA TAACAAAAGC TTATATGCTA TATAGATTGT   | 1620 |
| 15 | ATTACCCCTT GTTTTTAATT TTATTCATAA TTATTACAAA TATTTTAA TTAATCGTCA    | 1680 |
|    | TGCGTTACTT TCGTTCGTAC TCCTTTTATA AATGAACCAT GTAATATAAG CATGCTATTA  | 1740 |
|    | TCGACTCAGT TTGTCTAATG CTTCTTTGG TACTTCTTCC TTTTCAACTT CTTCAAAAGT   | 1800 |
| 20 | TTCTACATGA TGACCTTTAT GTGTAATTTT CAGTATCTA TGCGGTTTAA CATCAATGT    | 1860 |
|    | AGCAGTATAC ATTAATGTGG TCTCTTCCC TTCTTTATTA AAAACACTTT GACTATAACT   | 1920 |
|    | TCGGAGTTGA TCATCCATTC TAGTAGACAC CTCTGTGTT TTAACATAAG AATCATCCTT   | 1980 |
| 25 | CTTTACTAAT GGGTTAAATT GATCTGTTAT ACCATGGGAG TCTATGTTT TTAANAATGA   | 2040 |
|    | TAAAGCAGCA TAAGCGCCAA TGATAGTAAG TACAAGATAT GTKATTGTTT TTAANAATCGT | 2100 |
|    | TTCAAAAGA  | 2109 |

(2) INFORMATION FOR SEQ ID NO: 417:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 813 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 417:

|    |   |     |
|----|---|-----|
| 40 | GTTACATAAA TGAAAAACAA AAAGATAATT TTACTGCTTA TGCTACACCA GAACATAATT | 60  |
|    | ATCAATTTGG TGGTGTATG ATAGAAAGTG AAAAATTAAG CGAGTTACTA AAGCCAGCCA  | 120 |
| 45 | ATCAGTTAAA ATCACCAGAT GATATAAAAA AAGAATAAAA TAAAAAGAAG AGTCACTAAA | 180 |
|    | GTTAGGAGTT ACTTTAATGT CAAAAAACA TGTTTTTATA ATTATTGGTG TCATATTGTG  | 240 |
|    | TATATGTACA GTTTCTACGG TCATGCATTT TAAAAAGAAA TATGATGAAA AAGAAAAACA | 300 |
| 50 | AAAAGCGATT TACTACAAAG AACACAAGA ACGTATTACA CTCTATCTTA AGCATAAATAC | 360 |
|    | TAAAGAAACG AACACGATTA AATCTGTACA TTTACAAAC TTGGAAACAA GTCCTATGGG  | 420 |



# EP 0 786 519 A2

ATGCGCTGAA CATAATTATC AATTGTGGTG CGCTATGATA AAAAGTGAAG GAGTAGATAA 540  
 ATTATTAAAA CCAGCACATG AAAGAAAAATC ACCAGAAAAA ATCAAGAAG AATTAGATAA 600  
 5 AAAAGAAGGC CACTAGGGTC TTCITTATTT TTGATTTAAT CTTCACATAA TCTATGTCAT 660  
 TGCTATCGAA GGTGTTTCGC AATTAATATA AATCACTTCA TCATCACCAA TACTTCCCCA 720  
 10 GTTTGTGACA GTACATTAAAC ACAACGAAC CACGTTAATT TAAATGGAWT AaTAGTTTGG 780  
 CCATTATAAG ACAAATATAT ATCGanTAAC AAT 813

(2) INFORMATION FOR SEQ ID NO: 418:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 640 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 418:

AAAAGCAATC GTTGGTGGTG CTAATTCAT CGGCAACTCA TATGTAAGAG CTGGTCAAAA 60  
 25 TACACTTTAC AAAATGAGAT GGAATCCTGC ACATCCAGGA ACACACCAAT ATGCTACAGA 120  
 TGTAGATTGG GCTAACATCA ATGCTAAAAT CATCAAAGGC TACTATGATA AAATTGGCGA 180  
 AGTCGGCAAA TACTTCGACA TCCACAATA TAAATAAGCA ACATGAACAT AGGATCAAAA 240  
 30 GTCATCCCCC ACTATCAATC ATGGGGGATG ACCTTTGATC CCTTTTTTAT ACATACACAA 300  
 GCAAAAATAG CGGTGATTGT TTACCATCAA TTTTAACAAAT CACGCTACT TTGTCTTGTA 360  
 ATTCATGATT CAATTTTTGT TGTGTGCACA ACGACACTAA ATTATGTGTT TGCTATTGTC 420  
 35 GTGTTACAAC GATATGCGTC GTTGATTTAA CTTATCAAGT AATTGATTTA AATTGTCTAA 480  
 TTGACTTCC GATAACATT GACATCTTGC TTCAATCAAT TCGCAACGTG CATTATTAT 540  
 40 TTGTGAAATT AATGTACGTG CTTGATCAGT CAAAATTAAT TCTTTACATC TTAATCTTC 600  
 GCTAGATTGT TGACTTGTGA TGTACCCnT CAAAACATA 640

(2) INFORMATION FOR SEQ ID NO: 419:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1159 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 419:

EP 0 786 519 A2

|    |  |      |
|----|--|------|
|    | TTCTTGGATA CGTAATAAAT TACCTTTTGT CATAGGATTA AACATAGATC CAGGTAAAAAC | 120  |
|    | ATAAACATTT CTATTTTGCT CTAATGCAAA ATCTATCGTG ATGTGACTGC CACITTTGTTT | 180  |
| 5  | CTTAGCCTCA GTAATTAAAA CCCCTTTTGA CAAACCGCTG ATAATTCTAT TGCCTCAGG   | 240  |
|    | AAATCTATAT TTAGCGATTG GTGTATGTGG TGGATATTCA GATATACTA AACCTTTTTT   | 300  |
|    | TTCTATTTTA TTCTTAAATG CTAATGTACT TTGGGATAA TGTGTTGAT GGCCAAAGGC    | 360  |
| 10 | TAAAACTGCA ATTGTAGGGA GATTGTATTT TAAAGCTATT TGATGTGCA TTGCATCAGC   | 420  |
|    | TCCTTGAGCA AGGCCGGAAA CAATTGTAA ATATTTGCTT TTATCATTTG ATAATAAAAA   | 480  |
|    | TTCTAAAGAC TGTGGGTAT AACTGTAGA ATCTCTTGCA CCTACTACTG CCAATGATG     | 540  |
| 15 | CATATTATTT ATTAATTGTA TGTTCCTTTT ATAGAAAAGA AGTAATGGAA ATTGATATAT  | 600  |
|    | TTCTTTTAAT AGCACTGGGT ATTCAGAATC CATATACTTA ATGTAACCTA CTTTAAATTT  | 660  |
|    | CTGCAGTTCT TTAATAATAT CGTTATGATC AAGTTTATA AAAGCATAGT ATTTACGTAA   | 720  |
| 20 | TAAATGAACA TTTCTTCCC TATTCACCCA TTCACATAA TAACATCTIT TTTTCTTCC     | 780  |
|    | CTCCTCTTTA ATTACATTAG GATATGCCAT TAAAAATTGA TGAATTTGTT TAGTCGAAAA  | 840  |
| 25 | GTGTGCCAG TATAACTTAA GCAAAAATAG TTAAATCAAT AAATCAACTC CTTTTGTAA    | 900  |
|    | TCATACAAAA TCATATCTA TTTTGTGTTT ACATTTCTAA TACAAAAACA TTGTCGATGT   | 960  |
|    | AATGTATTTT TAAGGAGTAA AAATCTGAC TTA AAAAGTG AAAAGTATGT TGGAAAGAA   | 1020 |
| 30 | TTAAAATTC AATTCCAAC ATACTTTATA ATTAACCTT ATAAATAAGT TTGCAAAAT      | 1080 |
|    | TATATAGAAT TGGTCTTACT GGTGTGATGA AATCACCAT TAATCTTCA ACATGTGCAT    | 1140 |
|    | TAAAACCTTT TTTAAATTG   | 1159 |
| 35 | (2) INFORMATION FOR SEQ ID NO: 420:                                |      |
|    | (i) SEQUENCE CHARACTERISTICS:                                      |      |
|    | (A) LENGTH: 1879 base pairs  |      |
|    | (B) TYPE: nucleic acid   |      |
| 40 | (C) STRANDEDNESS: double   |      |
|    | (D) TOPOLOGY: linear   |      |
|    | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 420:                         |      |
| 45 | TCAATCAGCG ACTACAGAA GTGCGGTCT TCAACAATT GATGTGACAA CACTAAGTGA     | 60   |
|    | CCCCACTAAT ATTATCATGG GTATTTTAAT GTTTATAGGA TCTTCGCCAA GTTCGGTTGG  | 120  |
| 50 | TGGCGGTATT CGTACAACA CTTTCGTAT TTTAATTGTT TTTTAAATTA ACTTTAGTAA    | 180  |
|    | TAATGCCGAT AAAACATCCA TTAAAGTTTA CAATAGAGAA GTACACATTA TGGATATTCA  | 240  |

55

ATCAGCTACT GAAAAATGGTA AGCTTACATT TTACAAGTA TTTTGAAG TCATGCTGTC 360  
 ATTTGGAAC TGTGGACTAT CGCTGGGTGT CACAAGTGAT ATTAGTGATA TTTCTAAGGT 420  
 5 CGTACTAATG ATACTCATGT TTATAGGACG TGTGGCTTA ATATCAITTA TCATTATGAT 480  
 AGCAGGACGT CGAGAACCAG ATAAATTCCTA TTATCCAAA GAACGTATTC AAATAGGATA 540  
 ATATAATAGC AATCTAAGTT TAGTTAATGT AGATTTTAA CCGAACTTAG ATTGCTTTTT 600  
 10 TAGTTTGTAT TTTTAACTTA TTTTATAAGA CGATTGGTTT CGAAAATGGT AAAC TAGTAA 660  
 CAATGAGAGG TGTAAACATGA TGGAAAAAAA TGAAAACATT AATGTAGAGA TTTTAACTAC 720  
 15 GTCAGATATG CATAGTCATT TCTTAAATGG TGATTATGGT TCAATATATT ATAGAGCTGG 780  
 TACTTTATGT AACCAAGTAA GAGCACAAAA TCATCGCGTC ATTTTATTAG ATAGTGGCGG 840  
 AAGTTTAGCT GGCCTGTTAG CGGCCTATTA TTATGCTATT GTTGACACCTT ATAAACGACA 900  
 20 TCCAATGATA AAGTTAATGA ACAGAATGCA TTATGATGCT AGCGGTGTGA GTCCAAGTGA 960  
 ATTCAGTTT GGTATTATCAT TTTTAACTCG TTCATTTGCT TTGGCACGTT TTCCATGGTT 1020  
 ATCAGCAAA ATTGAATACA ATGTTACTAA GGAGCCTTAT TTTTCAACTC CATATTGTAT 1080  
 25 TAAACATTTT GGTGACTTAA AAATTGCTAT CGTAGGCGTC ACAGCAGATG GTTTAATGGA 1140  
 AAATGAGTAT TCTGAAATGG AGCAAGATGT ATCTATTGAA AAGCATTAG TGGCATCAAA 1200  
 ACGTTGGATT AGATATATCC ATGAAGTTGA AGAGCCAGAT TTTTGTATG TAAATTATCA 1260  
 30 TGGTGGATTG AATAAAATTA GTAATAGTAC GAAAAATAAA AAGGCAAGTT CGAATGAAGC 1320  
 TGAAAAATTA ATGGAAGAAC TCGGTGTTAT AGAATTAATG ATTACAGCTC ATCAGCATCA 1380  
 AACAAATAGA GTTCAAGATC ATGAAACGTA TTATGTTACG GCTGGTCAAG ATGCCAAAGA 1440  
 35 GCTTGATCAT CTTTCGATTA ATTTTAAAAA GAGAACAA CACTTATGATG TTGAAAGCAT 1500  
 TGATTCTAAA GTGATTGACT TAAATGAGTA TGAAGAGGAT CAAGAATTAT TAGATTTAAC 1560  
 ATTCTATGAT AGAAAAGCAG TGGCTTATTG GTCACAGGAA ATCATAAGTG ATAAAGGTTT 1620  
 40 GATGTTATCA GTAATGGGT TACAAGATT AGTCTGTCAA ACACATCCAT TTTCGCAATT 1680  
 ATTACATGAT GCAATTCCAC TTGCAITTTGA TAATGATATA ACATGTGTCC ACGTGCCTAT 1740  
 45 GAAACGGAGAG AAGGGGTTGA GTGACAGAT TCGAAATGAA GTTTTGTATC ATGCATACCC 1800  
 ATATCCAGAT AAGCCATGGG ATGACAATT aGTGGTCAAA ATATCAAGGT ATnTGTGGGT 1860  
 ATAGTTATTC ACCATTAGG 1879

(2) INFORMATION FOR SEQ ID NO: 421:

(1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 2710 base pairs

(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 421:

|    |  |      |
|----|--|------|
|    | AATTCCCCCT TTTTCCCAT CCATTThCC ATCCACTGGT CCTACGGGA CCATATTAT      | 60   |
| 10 | ThAAATTGGA hAAAAATTATc TTAAAGaTTA TwActACTCT TtAATCATTt TrGTGAATTA | 120  |
|    | AAAAAGTAG TGCAAAAGC AAAATATACT TTATACACTA CAATCATTt ATTTATAATA     | 180  |
|    | AAGTTTCACC AAAAAATGTT CCAACTAATG AAACGCTTG TTCAGCAGTA TGATTATTAC   | 240  |
| 15 | TGTCAATCAA TGGATTIACT TCAACTAAAT CCATTGAGGA AATTAAATGT GATTGATGCA  | 300  |
|    | GTAATTCCAA TGCAAAATGG CTTTCTCTAT AACTAAGAC ACCCAAACt CTAGTACCAG    | 360  |
|    | TGCCTGCGGT TTCAAGCGGA TCTAAAGCAT CAACACTTAA AGATAAATGA ACGCCATCAA  | 420  |
| 20 | CATTGCGTGA CTTCAAATAT TCTAATTGTAT TTTCaATTAC TTCCTTTATC CCCAATTAT  | 480  |
|    | CAATATCTGA CATAGTAAAT GTTTTAATAT TATGATCTTt GATAAATGTt CTTTCACCTT  | 540  |
|    | TATCTAAATC TCTCATACCA ATTAGTACGA TGrTTCTGG CTTGATTACA TTACnATTTA   | 600  |
| 25 | ATTCTAAAGT TTCTTTGGGA CCTTCGCTG TCAAAATCCT TAGAGGCATA CCATGAATAT   | 660  |
|    | TTCCACTTGG TGACTCTTCA GGTATATTTA AATCACCATG TGCATCATAC CAAATAACAC  | 720  |
|    | CTAAATTAIT ATAATGTTA CTTAATTGCTG ATACTGAACC TACCGCAATA GAATGATCTC  | 780  |
| 30 | CACCAAGAAC TAGAGGAAAT CTGTTATTTT CAATTGATGC TGAACCTCTt TtATTTAATT  | 840  |
|    | TTTGATTAAc ATCTATAATT TCATCATAAT TTCTTAATCC TTTTGTTCa CTATGAAATT   | 900  |
|    | TTTCAATGTT CACAGCAGGt ACCTTAATAT CCCCCTTGTC ATATACATCA AGGTCTAATT  | 960  |
| 35 | GCTTTAATCT TGAAATTAAT CCAGCATATC TAATTGCTGT TGGTCCTAAA TGAACACCTA  | 1020 |
|    | ATTTTCTTTG TCCAAATGTT GATGGTGCAC CTATAATATC AATTGCTTTT GTCTTTGTCA  | 1080 |
| 40 | TAAGCGTCCC CTTTGTCTCT ATGTAATTAA AGAATAATGT ATGCGCTTAC CATTATCAAG  | 1140 |
|    | CAATAGCTAC ACATATAATC TGTTTATCTT ATTACTTCAT AAAAAAGGT TCTTCATCTT   | 1200 |
|    | TTATGGTGGG AAGGTAAAC TTCCGTCTTT TTTTAATACA CAAAAAGCGC AATTGCCTCT   | 1260 |
| 45 | ATAATTTAAA GTGACCAAAC CCAACTAAA GGAGACAAGT GGCCTATGT GTAATGATAC    | 1320 |
|    | CTTAGAATTA CTAGAATAA AAGATGAAAA TATAAATAT ATAAACCAAG AAATTGAGGT    | 1380 |
|    | CATTATCAAA GGAaaaaaAG CAACAGTGGT TAATGCTGTA CTAACGTATA AGCCTTCGGC  | 1440 |
| 50 | CTGTTATTGT TGTGGAGTTA AAAATGAAGG ACAAAATCAT AAACATGGTA AGCGTGTTTC  | 1500 |
|    | TCGTATTACT TTACTTAAAA CTCAGGGSTA TAACACATAC CTCAACTTAG CTAAACAACG  | 1560 |

55



EP 0 786 519 A2

GGGTACTTT GATGAAATG ATTGAAGTGG ATAGAATTAA TGACGTATCT TGTGTATAAT 360  
 ACCTCGAAG TACTTTCATT GGAGGAAAA TAGTGACTCT ATTATTATT ATCGGGGTTTC 420  
 5 TCGTGCCAAT GGGTTATACC ATGCAGTTAA ATATTAAAA TGAACCTGTA ACAAAGCGCA 480  
 ATCTTTTAAT AACATTAGCT TTATCTACGT TAGGTATTTT AGTAACCGCG TTAGCAGGTG 540  
 TAATOGTTAC GAAACAAGCT TTTCTTTTAT TAAGTGTAGC AATTGGCTCA ATTTTACTG 600  
 10 GAATGCTTG GGGCCTTTTA CTAAGTGGTA GcTACGCGCT GATACGATT TTATCTAACG 660  
 CATTGGGCG TAAGTAATAG TAATCTGTTC ATCAAGTAGT ATCCGTGCTT GAAAACAAC 720  
 TAAACTCCT AATGTGGAAC TAGTTAAAA ATTCTAACT AGTACATTAG GAGTTTGTGTT 780  
 15 ATGCAGAATA AATTTAATTG TTAATTGAA AGTGCGGTAA AATCCACTA TTTATTGAA 840  
 AAAAACTGAG CAACCAATT AAACCTTGAT ATCCTAAGTA AATACATAAT AAAACAAGTC 900  
 CGACATAAAT TAAAAATCGC AAAATAGAAA GTCCAACCTT AAAAAGSAGG ATGACTAATA 960  
 20 GTGCTATTAA AATAATTGTT AGTATACTCA ATGTACAAA CCTCCTAATA CGCTTTTAAA 1020  
 TCCATAA 1027

(2) INFORMATION FOR SEQ ID NO: 423:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 569 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 423:

CGGCACATT TTAATTITATA CATATTTTAA AACTAAGTAA CAGTTTGAAG AAATCGTAGT 60  
 TCAATAATGT TAATTGTGAA AATGTATATA AACATAAAAA AATCATGTAT AATATATGTT 120  
 GTTAATTAAA CAGTTTCGAAA GCGAGATGAC ATTATGGGAC GTAAATGSAA TAACATTAAA 180  
 40 GAAAAAAGG CCGAAAAAGA TAAAAACACA AGTAGAATAT ATCGAAATT TGGTAAGGAG 240  
 ATTATGTTG CAGCAAAATC TGGTGAACCC AATCCAGAAT CTAACCAAGC TTTAAGGTTG 300  
 GTGCTTGAAC GCGCTAAGAC ATATTCAAGT CCGAATCATA TTATTGAAAA AGCAATAGAT 360  
 45 AAAGCTAAGG GTGCTGGAGA CGAAAACTTT GATCACCTAA GATATGAAGG ATTTGGCCCA 420  
 AGCGGAGCAA TGCTAATTGT TGATGCGTTA ACAAATAATG TAAATCTGAC TGCTCTGAT 480  
 50 GTGCGAGCTG CTTTGTGTAA AAACGGCGGT AATATGGGTG TATCTGATC AGTGTCTTAT 540  
 ATGTTGTCAT GTGCCACAT TTGGTATTG 569

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 3264 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 424:

|    |            |             |             |            |              |             |      |
|----|------------|-------------|-------------|------------|--------------|-------------|------|
| 10 | TGGACCTATA | AAATGGATGC  | ATCCAAATAT  | GGATGnGGGG | TanGGCGGGA   | AAATAGGagG  | 60   |
|    | TTATATAGTA | AGTCATTTAA  | GTGCAATAGC  | CATCCTTTTT | AAAATGAAAA   | AAATAGAAAG  | 120  |
|    | CTCAATAGTT | TGTTAAAGCC  | TTTAAATAGC  | GTGTTACGCG | AATTTTAGAA   | TGCTAAAAAAT | 180  |
| 15 | TGTCACAAAT | TTCAAAATTAC | GTGCTAAAGC  | TTGAGATATC | AATATTATT    | TGGCGATAGAG | 240  |
|    | TGTAATTTGA | CTCCGCTTAAT | ATCTCAGTTT  | TATGTTTGTG | ATTTGGTAGC   | ATAATATTAT  | 300  |
|    | AATAAAATAA | AATTGGTTAC  | CTTTAAITTC  | ATAGATAGTA | TTTTTACGTG   | TAGTCACGTTG | 360  |
| 20 | TAAAAATAAT | TCAATTAGST  | TAGAGACAT   | AATATAGAA  | ACATTTCGTG   | AAAJAGTTGT  | 420  |
|    | ATTAATCGGA | GATGGATCTG  | TAGGATCAAG  | CTATGCGCTT | GCAATTGGTT   | CGCAAGGTGT  | 480  |
|    | TGCTGATGAA | TTTGTAATTA  | TTGACATTGC  | AAAAGACAAA | GTA AAAAGCAG | ATGTTCAAGA  | 540  |
| 25 | TTTAAACCAT | GGTACAGTCC  | ACAGTCTTTC  | ACCAGTTGAT | GTGAAAGCAG   | GTGAATACGA  | 600  |
|    | aGaCTGTAAA | GATGCAAGATT | TAGTGTGTTAT | TACAGCTGST | GCACCTCAAA   | AGCCAGGTGA  | 660  |
|    | AACACGTTTA | CAATTAGTGT  | AAAAAAATAC  | TAAGATTATG | AAGAGCATCG   | TTAAGAGTGT  | 720  |
| 30 | TATGGATAGT | GGCTTTGATG  | GATATTCTTT  | AATCGCGGCA | AACCTGTAG    | ACATTTTAAC  | 780  |
|    | AGATTTTGTG | AAAGATATGA  | CTGGATTACC  | AGCAGACGCT | GTTATCGGTT   | CAGGTACTGT  | 840  |
| 35 | ATTGGACAGT | GCACGTTTAC  | ATATTTTAAT  | TAGCAAAAGT | CTTGGTGTGT   | CACCTTCAAG  | 900  |
|    | TGTTGACGCT | AGTATTATTG  | CGCAGCATGG  | TGATACTGAA | TCTGCAGTTT   | GGTCAACAAG  | 960  |
|    | AAATGTAGCA | GGTATTTTCA  | TATATGACAC  | ATTTAAAGAA | CAAACCTGTA   | GCGAAGcTAA  | 1020 |
| 40 | AGCGGAAGAA | ATTTATGTGA  | ATACACGTGA  | CGCTGCTTAT | GA AATTATCC  | AAGCTAAAGG  | 1080 |
|    | GTCAACATAC | TATGGTATTG  | CATTAGCATT  | GATGCGCATT | TCAAAGCCCA   | TTTTAAATAA  | 1140 |
|    | TGAAAATAAT | GTCTTAAGAT  | TTTCTATACA  | ATTAGATGTT | CAATATGGTG   | GTCACAAAGG  | 1200 |
| 45 | CGTTTACCTA | GGTGTAACAA  | CATTAGTTAA  | CCAACATGGC | GCAGTTAAAA   | TTTATGAAAT  | 1260 |
|    | GCCATTAAGT | GCCGAAGAAC  | AAGCGTTGTT  | CGATAAATCT | GTTAAAAACT   | TAGAAGATAC  | 1320 |
|    | ATTTGATTTA | ATTTAAATAT  | TATTAGAGA   | CTAGGCTTAT | TTTAAGTATT   | AATTAGAAAT  | 1380 |
| 50 | ATATTAAATG | TAGAGAGATC  | TATTAGTGT   | CGACGTAACG | GTGGCACTGA   | TAGGTCACAT  | 1440 |

55

EP 0 786 519 A2

|    |            |            |            |            |            |            |      |
|----|------------|------------|------------|------------|------------|------------|------|
|    | TATGACTCGT | GAAAATTAG  | TGACATGAAT | AAAAATGTTG | AGGCGTCATT | GAGTATAAAG | 1560 |
|    | GAAAGTAGTT | CTGCATTAA  | CACGAAGTAG | AGCATGACMA | CGAGGAATAA | CTATAGGGAG | 1620 |
| 5  | ATGGTTTITG | AATGACGATG | TCTTGTATCA | ACATGGTACA | TTAGGTACGT | TAATGGCTGG | 1680 |
|    | CTTACTAGAA | GGCAGAGCTA | CAATTAATGA | ATTGTTAGAA | CATGGGAATT | TAGGGAATTG | 1740 |
|    | AACGTTAACA | GGTCTGATG  | GGGAATAATA | TTTTTAGACG | GAAAGGCATA | TCATGCTAAC | 1800 |
| 10 | GAGCATAAAG | AATTTATAGA | ATTAAAGGCG | GATGAGAAAG | TACCGTATGC | ATCGATTACT | 1860 |
|    | AATTTTAAAG | CGAGTAAGAC | ATTCCATTG  | CAACAATTAT | CACAAGATGA | TGTATTGTC  | 1920 |
|    | CAAAATMAAA | ATGAAATGTT | AAGTGAGAAT | TTATTTTCGG | CTGTTAAAT  | TTATGGCACA | 1980 |
| 15 | TTTAAACATA | TGCATGTACG | AATGATGCCT | GCTCAGCAAC | CGCCATATAC | ACGTTTGATT | 2040 |
|    | GATTCAGCAC | GCAGACAACC | TGAGGAAAAA | AGACAAGATA | TTGCTGGTGC | CATTGTTGGA | 2100 |
| 20 | TTTTTTACAC | CAGAAITATT | TCATGGCGTA | GGGCTGCTG  | GTTTTCATAT | ACATTTTCGG | 2160 |
|    | GATGATGAAA | GAGCTTATGG | TGGACATGTT | CTTGACTTTG | AAGTGGATGA | CGTTGTCGTT | 2220 |
|    | GAGATACAAA | ACTTTGAAAC | ATTCCAACAA | CATTTCCCGG | TAAATAACGA | GACGTTTGT  | 2280 |
| 25 | AAAGCGAAAA | TAGACTATAA | AGATGTGGCA | GAAGAAATTA | GAGAAGCTGA | ATAGTCTAAT | 2340 |
|    | ATAATTAAAA | GACCTTAGCG | ATAITAGGAA | CAGGTGGTTC | TAAATGCATC | GTTAAGGTCT | 2400 |
|    | TTTTATATTA | TATATGTGCT | TACATATTTT | GTTGATACGC | GGTAAAAAG  | CTATAAATAG | 2460 |
| 30 | TATCAAGATA | TGGGGTGTC  | AGAATGTGTT | CACGTGCACG | TGATAAATA  | AAGCCTTGAA | 2520 |
|    | TGCGTTCGAC | TTCAAGAGGT | TGCTGATGCG | CAATGTCGTA | ATACATGCTC | GTTCCCATTT | 2580 |
|    | CGTCGGGATA | ACCTTGATAT | ATAGTCATAA | TAGTATCGAC | AGTTTGTTC  | GAAGAGTTTA | 2640 |
| 35 | ACCCCTCTGC | CTGTGCAACA | CGACAACCAT | CTAGTAATAG | GTGCTACAC  | AGTGTACGTA | 2700 |
|    | TTTCAGGATT | ATGCATGATT | GCAACGGTCT | GTCTACCGAG | TGCTGTGATA | GAATTAATGC | 2760 |
|    | CCAAGTTAAC | TAACAATTTA | TACCAAATAG | CTTGTGGAAT | ATTGCTTCT  | AATACAATGT | 2820 |
| 40 | CGATTTGACT | GTCTTGACC  | AAATCTCTAA | ATTGTCGAGT | TAATGCATTA | TCTTGATACA | 2880 |
|    | GTAATTGATA | ATCTCTGAAG | TGCGTAACAA | CATCGCCTTT | CTTTTGACCA | CTTATATAGA | 2940 |
|    | CAACTGCTTG | GCATACGTTT | TTAAATGAAA | ATGTTGCGAG | TTGACCATAA | CCATTTTGGG | 3000 |
| 45 | CTAAAATGAT | GAGCGTGTC  | TCATGTGCAA | GATGAGTTAA | ATGAGGAATA | ACTGCATCAA | 3060 |
|    | GTTGATGTGT | TTTGACTGCA | ATAATAATAA | CATCAAATGT | GTTTGTGACA | TCTTCATAAC | 3120 |
| 50 | CTTTCACAAC | AATATCTTGT | GCAGGTGCAT | GTGGTACAGT | ATAATATGTT | ATTGTTTGGG | 3180 |
|    | CGTGCTCCCG | GATAAGCGTT | GTATGAGGCA | ATGATTGTTG | TAATTCATAG | GCAATAGTGT | 3240 |

55



## (2) INFORMATION FOR SEQ ID NO: 425:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 796 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 425:

```

AGAACGAACC TTTAATACGT ACGTGTGAC CACGAACGT CATGACAAAT TGATAAAACG      60
AATAGTTTTT CATTAGTTCA TTGTCACATC AATCACTTTT GTATCACCTT TAATCACAGC      120
ATTTCATCA TAAATATTAA TTGAAGCTGC TTGATCAGTG TTAGTAATTA TAATTGGTGA      180
AATTACAGAT TTAGCGTTAT TATTAATATA TTCAAGGTTG AATCTTACTA ATGGATCTCC      240
GACGTTAACT TCGTCACCAC TAGACACTAA CACTTCAAAT CCTTCACCGT CTAATTGAAC      300
TGTTGCTAAA CGATATGAA CTAATAATTC TAATCCGTTA TCTGCTTTTA ACCCAATCGC      360
ATGCTTAGTT GGAAGACAT TGTCAACAG TCTGCAATT GGAGACACAA CTTCCTCTTC      420
AGTTGGATTA ATACAAAAC CTTCGCCCAT CATTTTTGT GCGAATACAG GATCTGGAAT      480
ATCTTCAAAT TTCACGAATT CTCCAGTTAA TGGTGCATAA ATTGCGATAT CTTTCTGAAC      540
TTCTTTGCCT TTTCGGAATA ATTTTTTAAA CATACTTCC ACTCCTACTT ATCAAAATGT      600
GATATTAAAT CGCCATAACC CAATCTTCT AACTTTTCAT ATGGAATAAA TTGAATGCA      660
GCGGAATIGA TACAGTATCT TAAGCCGCCA CTTTCTTAG GTCCATCATT AAAGACATGT      720
CCTAAATGAC TATTGATTC TTCTGAACGC ACTTCAGTTC TCAACATACC AAATGATTGT      780
TCGACTAATT CTATAA                                     796

```

## (2) INFORMATION FOR SEQ ID NO: 426:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 3393 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 426:

```

ATCCAGCAAA TCTTCTTAT CAGGTTCTGT AATTTTCTTA GTATCCATCT TGATTAGCTT      60
TGATAACTTT TCAGCCGTAT CCAACATTC CGATTGTGTT GTTTTGCAC CCCTAGTATA      120
TGTAATAGCC ATTTTAGAAG CATTATCAAC TAAACCTTTC CCATTTCTGT CTAATAATACG      180

```

|    |  |      |
|----|--|------|
|    | ATGTGAGCCT TGTGCGATTT GTAAATAACC TAAACGTAGT ACTAGTACTG CAAAAATAAA  | 300  |
|    | TACAATCACA CCAAATATAA AGTTAATICT CTGTGTAATT GTATTTTGAA CGATTTCATC  | 360  |
| 5  | ATTTGATTTT TCTTTTAGTC TTTTAAACAA AACTACCTAC CTCTATTCAA AGCTTTTCAC  | 420  |
|    | TTTAAATCAT ATATGAATTT AGAAATATT TCTATCTTTT TGACAAAAA ATAACGGTCT    | 480  |
|    | CATTAAAGAG ACCGAACAG TAATCATACT TTATTTTGTG GCATTATATA ATTCGTCAAC   | 540  |
| 10 | TTTTCCCGAG TTAACATACAT TCCAAAATGC GCCAATGTAG TCAGGCGGCT TGTTTTGATA | 600  |
|    | TTTTAGGTAA TAAGCGTGT CCCATACGTC TAAACCTAAA ATAGGTGTTT TACCCTCAGT   | 660  |
|    | TAATGGATTA TCTTGGTTTG GTGTAGTCAC AATTTCTAAC TGGCCATTGT TTACGACTAA  | 720  |
| 15 | CCAAGCCCAA CCTGAACCAA AGCGTGCAGC TGCTTTGTCA GCAAAATCTT TTTTAAATTC  | 780  |
|    | TTCTAAAGAA CCCCATGTGT CTTTAATTTT TTCTACTACA GTACCTTTT CTCTGAGTT    | 840  |
|    | TGTTGAAAGT AACTCCGAGA ATATGAATG GTTTAAATGT CCACCGCAT TATTACGTAC    | 900  |
|    | AGCAGTTTGG ATGTTAGCTG GTACACTGTC TAAATTAGCA ACAATTTCTT CAATAGATTT  | 960  |
|    | AGATTTCTAA TCTGTACCTT CTACTGCAGC ATTTAATTC GTACATACG TGTATGATG     | 1020 |
| 25 | TCTGTCATGG TGAATTTCCA TAGTTCTTT GTCAAAATGT GGTCTAATG CATCAATGC     | 1080 |
|    | GTATGGTAAT TTGGTAATT CAAAGCCAT AAATAATCAT CCTCCTAAAA TGTCTGAAG     | 1140 |
|    | TAAATAATAA CAAGCAGTGA CTGGTTCAAC AAAGAATTTG CTTAAATCTT ACTACTTATT  | 1200 |
| 30 | ATTTTCTCTA CTCATTTAAT ATAACTCAA TCAAAAATAA TTAACATTT TGTATATAAA    | 1260 |
|    | AAGTTAACAG ATTTGCCATA AAATCATACG AACGGAGTAT GAAATGAACC TTTATCTTCT  | 1320 |
|    | ATAATTTAAA AAATGAGATT TATGCATACA TCGGACCAA TGTGCATAAA TCTCATTTCT   | 1380 |
| 35 | TTATATTAAT CTGGCAAGA CTCACATACA CCATAAATC CAAGTTTGTG TTGTGGAATA    | 1440 |
|    | TTAACCACAG GTAGTGATAA TTTTATCTGA TCTATTGGAC AATAATCTAT TACCTTTGTA  | 1500 |
|    | TCTCCACACT TTTCACAGAT AAAATGATGA TGATGATGGT TTGTACAAGC GATTCTAAAC  | 1560 |
| 40 | TTCAATTCAC CATCAAGTTC TGTATTTTCA ATAATTCCTA AATCTTTAAA TAAGTGCAGG  | 1620 |
|    | TTTCTATATA TTGTGTCGAA TGAAATTTCA GGATAATTTT CATCCATAAC TTGTGTATA   | 1680 |
|    | TACTTTGCGT TTATATACTT ATCTTCTTCG AAAAAATAT CTAACATATC TTTACGTTTA   | 1740 |
| 45 | TCTGTATATT TTAACCGTT CTCTTTTAAA ATTTTAATAG CATCATTTGT ATTCATTGAT   | 1800 |
|    | ATTAGCTCCC TTTTAAACT TCATTGCGAT TTTCTGATAA GCCATTGTAA TCATAAGTAA   | 1860 |
| 50 | AATAACAAGT AGAATAACAA TTACACCACC CGGAGAAATG TCCATATAGA AAGTAGGAC   | 1920 |
|    | TAAGCCTAAT ATTACTGATA ATTCACCTAA AAATACACT AGTAATATCA ATTGCTTAAA   | 1980 |

TATCCCTACA ACACGCATG AGGCAGAAAT AACCATCGCT ACAATAACAA TAAATAAAAA 2100  
 TTGAATCCAT TTAGGAATGC CAATGACTTT ACTATATCCC TCATCAAATG ACAATATAAA 2160  
 5 TAATTCTTTA TAAACAATG TAATAACAG AACAACTATG ATGGCAATGA CAATAATCGT 2220  
 TGTAAATCA CITATATTCA CTGCGCTTAT TGAGCCAAAT AGCAATCCAA CAATTTCTTG 2280  
 ATTGAACCCA TCAGCTAATG AATGAAGAT TGCACTCAAG GCGATACCAG CACTCATTAT 2340  
 10 AATTGGAATA GCAATTTCTT GGTAAAGCAGT GTATGACGTT CTAAATTTTT CAATTAGAAG 2400  
 GGCACCTACT ATTGCGAATA AGATTCCAAA CCACATTGGA TTAATAATA CTAGTGTGG 2460  
 CATAATAGTA AGTAAAACA TACCGAAAGA TATACCACCT AAAGTTACAT GACTTAGAGC 2520  
 15 ATCAGCTWTA AGTGATAGTC GTCTAACAA GATAAAAGCA CCGATTAGAG GCGCAATAAA 2580  
 ACCTATCAAG ATACCACTAA TTAAGAGTA CCTCATAAAA TCAAAATTCA ATAATGCATC 2640  
 TATCAATTGT GACACGCGCT TCCATTTTAA ATAAACTCAA ATCTTTATTA ATTACAACAT 2700  
 TCTCGATTAT GCTGATGATC GACAAAACGT ACAGGATGTC CATAAATTTT TGAATTTTCA 2760  
 ACITTCATCA GTGATTTAAA CTCATCAGTT GTACCATGGA AATGCAATG CTTATTTTAA 2820  
 25 CATGCTACTT CAGTAGCAGT ATCTGCTACA ACACCGATAT CATGAGTAAC TAAGATAATG 2880  
 GTGATACCTT CTTGTTTTAA TTGATCTAAA GTATTATAAA ATTCACTTAC ATGTTTTGCA 2940  
 TCAATACCAT TCGTTGGTTC ATCAAGTACT AATACTGcAG GTTCTGAAAT CAATGCTCGA 3000  
 GCAATCATTa CAGTTGTTG TTGACCACCT GATAATTCTG CTATATTTTT ATGAATTAAA 3060  
 30 TCACTTATAT TCAGTCTTTC TAGTACTTTA ATCACTTTTT CATTATCTTT GCTATTAAAT 3120  
 GTTTGGAAAA GACGTTTTGT CTTTGTTAAT CCGCTTAAAA CAACTTCTTT AACACTTGCT 3180  
 GGGAAACCTG AATTAAGGCG ATTTGCTTTT TGTGATACAT AGCTTAAATT AATTGATGTT 3240  
 TcTtATTTTT AAAATCAATA CCTTCAACAA AAATCTCACC ACTTTGTAAA GGTAAATAAC 3300  
 CTAGAATCAA CTTCAATAAT GkTGATTAC CAGCACCAT TGGKCCAACA AwTGCTAAAA 3360  
 40 ATTCACCTTT AITTATTTTG AATGnnnTAT ATT 3393

(2) INFORMATION FOR SEQ ID NO: 427:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1123 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 427:

CCTGTTTCT TAGCCTTTTC GTATATGCGT TCATAAGAAG AATTCTTTT GATTTCAGAT 120  
 TGTAAAGCAC TGTTTTCACT TGATTGTTA TCTATTTTAT AATCTAAATC TGCAATCTTT 180  
 CCTCGGTAT CATACGCATC CATTTTTAAA GATAGCATAT AAATACCTAA CATAGCAATT 240  
 ACAGTAATCA AAGTTATGTA TAAAACCTTT TCAAAATTAG TTAATTGTAC AACCACCTTT 300  
 CTGAAAACAG TCTTCTTTTC GGGCTTAGTT TGTGGTTGTT GCTTCGGTAT ACTATTATAA 360  
 ACTTGTTTGT CATATGGTTG GTACACTTTT TCTACAGCCA TTATAAAATTG CTCCTTAATT 420  
 AAGTATTTCA GCTACACGTA ATTTGCGGCT TCGTGCTCTG TTAATGTCAT CTAATCTTTC 480  
 TTCTGTAGCG GTAATCGGTT TACGATTAAAC ACGCTTTAAC TTAGTGATAT ATGCTTCTGG 540  
 TATAACTGGT AATCCTCTTG GTACCTCTGG ACCTTTTICA TATTCTTGGA ACACCTGTTT 600  
 ACATAAACGA TCTTCTAAAG AATGGAAGT GATTACCGAA ATCTCGCCAT CTACTTTCAC 660  
 TAATCAATC GCTTGTTCTA TTGAATCTTC AAAAGCTGAC AATTCTGCTT TACTGCAAT 720  
 TCGTAGTGCT TGAATACTC GTTTTGCAGG ATGTCCGCCT TTCTTCTTG CTTTTCAGG 780  
 AATACCTTCT TTTATAATGT CAACTAATTC TAATGTTGTT GTTATTGGTT GTTGTTCGG 840  
 ATGTGCTTCG ATTCTTCGAG CTATCTGTTT TGAATAATTC TCCTCGCCAT AGCGATAAAA 900  
 AATCTTCACT AACGCTTCAT ATGACCAATT GTTAACAATT TCATATGCTG TTAGTTCTTG 960  
 TGTTGGTCC ATACGCATGT CTAATGTTGC GTCATGGTGA TAACTGAATC CTCGTTCTGG 1020  
 AATGTCGAGT TGTGGGCTTG AAACACCCAA GTCGTAATAA ATTCCATCTA CTTTTCAAT 1080  
 GTTTAAGTCT TTTAATATTT GAGTTAAATC ACGGAAGTTG CTA 1123

(2) INFORMATION FOR SEQ ID NO: 428:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 838 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 428:

TATATCTCC AnACCATAAA AATGGAATTA TAGTGAAGAC CGCTATGATC CAGGATGACC 60  
 ACGTATTATA GAATCTTTG ATAGAGAAAC GAATCaACGT ATCCgACATC aCGAAATGAA 120  
 ATTAGAAGAT TATAAGATG AGTTAAGAAG AGAATATCTA AAACAATCTG ACAGAATTGA 180  
 AGGAGATGAA TAAGCGTGGG ACTTGATTTT AGTGGTTTAC CAGATTTAGC AGTATTGGAA 240  
 CAAATGAAGG AAAAAGAACA GATTAGTGAG GTTATTGCGC CTGAACATGT TCGTATGCAT 300

|    |            |            |             |             |            |            |     |
|----|------------|------------|-------------|-------------|------------|------------|-----|
|    | CATTTCAAAA | AATTTGAAGA | TGATTTTAAA  | AATGGGCGAC  | AAGGGGCTTG | GGTAAAAAAT | 420 |
|    | GCCACAGAGC | AATTAAAAAG | TATTAGTAT   | GATTTAGAAA  | AAATTCAGAA | TATTAAAGTA | 480 |
| 5  | TAAAAAGGTA | TTAAGAAAAA | CAATAAGTAT  | ATAATCCATT  | TAATAATAAA | TGAATTATAT | 540 |
|    | AGTTCATAAT | TTCGACTATA | AGTGCGTATT  | AGCCATTATT  | TTTCGGGATC | TATGTCAAAT | 600 |
|    | CGGACTAATG | AATTCAATTA | TGGAAAGTTAA | GCAACCAATC  | TTTGTTTAA  | TTCTTTTTTT | 660 |
| 10 | TTTTTTGGAA | ATAAAGATT  | TGAACAATAT  | AAAAATTAGAT | TATGTTTTTA | CGAATTTTGA | 720 |
|    | TGTTTCTGTA | ACTATATATC | ATCTAGTCGT  | TATTCATT    | TGTAATAAT  | GACTTCAAAC | 780 |
| 15 | TGKATGAAGG | TGATGGGGAT | TAAAGGCGTC  | ATCGTAGGT   | TCTAAAGAAG | TAGANTTTT  | 838 |

(A) LENGTH: 1150 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

|            |            |            |             |             |             |     |
|------------|------------|------------|-------------|-------------|-------------|-----|
| NGTGTTCAAA | TAATAGAATG | GCTAATACCC | CAACAATAAA  | TAGGAATAAA  | TATGTTCATG  | 60  |
| TCGTAATTTT | AGTAAATAAT | GAAATCTAA  | AATTAGTAAC  | CCTATTTTGA  | ATATAAGCTT  | 120 |
| TAACCTCTAA | TAATACTGGG | AAGCGGATTG | ATCCTAAAAAT | TATTAAAAAT  | ATACTATCG   | 180 |
| CTTGACAAA  | ATAATCATGT | GCATAAGGAA | TAAATGACTT  | ACCTGTAATA  | TCTAATCCAC  | 240 |
| CATTGGTAGT | GGCAGAAATA | GAAACAAAAA | CACCTTGCAT  | GATTGCAAT   | TTTAAATCTG  | 300 |
| GATTATCTCG | ATAAAGAGTA | AATGCTAATA | ACATGGCTCC  | TACGAGTTCC  | ATTACAAAATA | 360 |
| TTGATTTTAC | AATATCAATA | ATCAATTTGA | CGGTACCACT  | CATTGTGTTT  | TTGTTATTAT  | 420 |
| CTAACATAAT | TAAGTACGCT | TCTCTAATTC | CAATATGTTT  | ACCTAGTACC  | ACCCATAACA  | 480 |
| TCGTACCAAT | TGCCATGACG | CCAATTCCAC | CAATATTTAA  | TATCAGGAGG  | ATAATTAATT  | 540 |
| GTCCAAATGT | AGAATAGGTA | TCGACAATAC | TTATCGGAGA  | CAATCCAGTA  | ACACTAATTC  | 600 |
| CGGATACCGC | AACAATAAAT | GTGTCAATTG | GATTTACTTC  | TACACCTGGT  | TTATGAACAT  | 660 |
| ACGGTAAGTT | TAATAATAAA | AATGCAATTA | CAATTGCGAC  | GATATAGTAC  | AATACAATAC  | 720 |
| CTTGCTAGGT | GCTTGATCTT | TTTAAAAAAT | GGCTAAAAAT  | TSACACGGTTG | TTCACTCTCA  | 780 |
| CATTATTTTA | GTTTAATATT | ATTCTCTTTC | TCTTTTACAT  | TCTTATAAAT  | CTTCGCTGTA  | 840 |
| ATTGATTTC  | AGTCACTTCT | ATGACTAAAT | ATAATCTGCC  | TAAAGCGTAA  | ATCATCTTCT  | 900 |

AAACCATTTG TGTCAACTTG ATCTACAACA ACACCGTTCT TAACTTTTCC TGGCAATTTA 1020  
 ACTGCTTGTC TTTCAAAAC ATTTAGACTG ACAATATTCT TCATTTTAAC ACCTACATCG 1080  
 5 GGATAGTCAA TTTTACCTTT tGTTTCTAAA TCITTTTACAA TCITTTGTAC TTCATTAAaCA 1140  
 GGtATTGCAA 1150

## (2) INFORMATION FOR SEQ ID NO: 430:

10 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 797 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 15 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 430:

20 AAGAATATTT AGGATTTTAT GTTTCGCAAC ACCCAGTAGA TAAAAAGTTT GTTGCCAAAC 60  
 AATATTTAAC GATATTTAAA TTGAGTAACG CGCAGAATTA TAAACCTATA TTAGTACAGT 120  
 TTGATAAAGT TAAACAAATT CGAACTAAAA ATGGTCAAAA TATGGCATTG GTCACATTAA 180  
 25 ATGATGGCAT TGAACCTTTA GATGGTGTGA TTTTCCCTAA TCAGTTTAAA AAGTACGAAG 240  
 AGTTGTTATC ACATAATGAC TTGTTTATAG TTAGCGGAA ATTTGACCAT AGAAAGCAAC 300  
 AACGTCAACT AATTATAAAT GAGATTGAGA CATTAGCCAC TTTTGAAGAA CAAAAATTAG 360  
 30 CATTTGCCAA ACAAAATATA ATTAGAAATA AATCACAAT AGATATGTTT GAAGAGATGA 420  
 TTAAGCTAC GAAAGAGAAT GCTAATGATG TTGTTTATC CTITTATGAT GAAACGATTA 480  
 35 AACAAATGAC TACTTTAGGC TATATTAAAT AAAAAGATAG TATGTTTAAT AATTTTATAC 540  
 AATCCTTTAA CCTAGTGAT ATTAGGCTTA TATACTTTT ATCAACTAAT AAATTATGAT 600  
 ATAGTAAACT GATGGTTAGA TATTTyTAA CCATCAGTTT GCGTktATAT TAGTTTTTTA 660  
 40 TGCTTATTAT TTTTATGAGT TTCACCTTAC ATTATTGATT AATCCATAAG AATAATTAGC 720  
 ACAAAAAGCA GTATACATAA ATTGAGTAAA GAATTTTGTC GATATAACCG nGCGGAAAAA 780  
 TAATAACnTT TGGATn 797

## (2) INFORMATION FOR SEQ ID NO: 431:

50 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1466 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 55 (D) TOPOLOGY: linear

|    |  |      |
|----|--|------|
|    | TTTCCAAGAG AAGTTGCGGA AGTAATTAAT AAAACGCATC ATAATAAATT GGTCAITTCG  | 60   |
|    | ATGATTTCGK CACAAATCGA TCCGATAGA ATGGATTATT TACAACGTGA TCCGTAITTC   | 120  |
| 5  | ACAGGTGTAT CATATGGTGC TTTTGATATG GAGCGTATT TAAGATTAAT GCGACCTTCT   | 180  |
|    | AAAGATGAAG TACTAATCAA AGAAAGTGGT ATGCATGCAG TTGAAACTT TATTATGAGT   | 240  |
|    | CGTTATCAAA TGTATTGGCA AATTACTTTC CACCCAGTTA GTCGTGGTGG AGAAGTGCTG  | 300  |
| 10 | CTTAATAATT GATTGAAACG CGCAAAACAG CTTTATAATG AAGGCTATGA AITTAAGTTG  | 360  |
|    | CATCCACATG ATTTTATTCC ATTTTGTGAA GAGACAGTTA CGATTGAACA ATATGTTGAA  | 420  |
|    | CTCGATGAAG CGGTAGTTAC GTATTATTTG GAAAAATGGA CAAAAGAAGA TGATGCTATT  | 480  |
| 15 | TTAAGTGATT TAGCAAGTCG ATTTATTAAT CGAGACTTAT TTAAATATAT TCCATTTGAT  | 540  |
|    | GGCTCAATTA TTACAATATC AGAAGTCAA GAACGTGTTG AAGCAGGTGG TATTAAATCCA  | 600  |
| 20 | GATTATTATT TTGTGAGTGA AGCATTTTCT GATTTGCCAT ATGACTATGA TCGACCGGGG  | 660  |
|    | TCAAATCGCA AACCAGTTC A TTTATTAGA CAAGATGCTA CGATTAGAGA AATAAGCAAT  | 720  |
|    | CAATCATTAG TCATTCATAG TATTACAGGC ATTAATCGCC AAGACTATAA ATTATATTAT  | 780  |
| 25 | CCTAGAGAAA TGGTTGCAAA GATTAAAGAT AAGACAATTA GAGAAGCTAT TGAAAATTTG  | 840  |
|    | ATTAATGAGC TTAATTAAAC AGGGCTAAAA TTGTTATCGT TAAATATGGA GGTATATCA   | 900  |
| 30 | TTGTCGTAGA AAAAAGGCTT TAATTTTAAT ATCATAAAAA ATGACCCCTCT AGATGGTCAT | 960  |
|    | AAAGGTACAA ATATTGGTTC AATTAGCTTA GACAATATTG CACCAGTTTT TATCGATGTT  | 1020 |
|    | GCTAACAAAG AAGCATTTAT TGATATTGGA GGCATGCATG CTCGTGCCAA AGTTGAAAAA  | 1080 |
| 35 | GGTGTGAAAT GGATTACTGA TAAAGCTGCT GTTGAAGGCG ATGAAGCTAA AGAATATTGG  | 1140 |
|    | TTGTGTTGGG TAACAACAGA ACSTAATGAA CAAGGACCAT ATTACGCTGG TTAAACAGCG  | 1200 |
|    | TGCTATTAT TAGTGAATAA AGCAATTCGT CGTGGTTATA AAAGTATGCC TGAACATGTT   | 1260 |
| 40 | AATATGATGG ATAAATCAAT GAAACATCAT ATTATCATAG ATCAAATTGG TGAACAGAAT  | 1320 |
|    | AAAGCTATIT TAAAGACTIT TTAAATGAAC CATGATGAAG GTATGTGGAA GCATTCTTCT  | 1380 |
|    | GATGCTTTAC ATCAAGCATT TAAITAAATA TTAGAACTA AAATTTCCCA ATTAATCTAT   | 1440 |
| 45 | AAAGATATGA TTCATTCTCT AATGAC                                       | 1466 |

(2) INFORMATION FOR SEQ ID NO: 432:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 2304 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

# EP 0 786 519 A2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 432:

|    |             |            |            |            |             |            |      |
|----|-------------|------------|------------|------------|-------------|------------|------|
|    | TTGTACGAAT  | ACAAAGTTGC | ATTAATATAG | AATCATTAT  | GCTTTTAAAT  | ATCACAAAAC | 60   |
| 5  | AAATGGATTA  | AACTAGCTAA | AATAGTCAAA | ATTGGAATCA | AGATTACATA  | AATATCATAA | 120  |
|    | CCTCTAGATA  | TTAGTATACA | TAATATAATT | AGCGACCTTA | TTATAATTGC  | TAAATTATAA | 180  |
|    | GCCAAACTAA  | TTTTATATTT | CATTTCAAAT | CCCACCTTAA | TAGCATTTAT  | TAAAATTCTA | 240  |
| 10 | TAATGTTTAC  | ATAGACTTCT | TAACAAACTT | TAACCCAGCT | AAGCTAAGTA  | TTGAAATAAT | 300  |
|    | TACACCTCCT  | AATATAAATA | TATTACTTAA | GCTTAGTAAA | GGTATAATTA  | TACTTATCAG | 360  |
|    | TCCTAAAGAC  | AATGTATCCG | CTGcATAAAT | CGaTGTAGAT | GAGATACTAA  | ATACTTCCC  | 420  |
| 15 | CATCAAAATGA | TTTGGCGTTT | TTATTGGAAT | GGCAACTGAT | CTTGTTAGTC  | CCTCTATAGA | 480  |
|    | TTGTCCAAGT  | CCCAATAATG | TTGCACCTAT | ATATAATATC | GCCACACTTG  | GAACACATT  | 540  |
| 20 | AATAATCGAT  | AAGCCAATTC | CCCAAACTAA | AACACCAATA | CTAAATTTAA  | AGATTAATCG | 600  |
|    | CTTTCTGAC   | AGCAAAACCA | TAATCAATGA | CATTAATAAA | GATGCTATAC  | CTAAACATGA | 660  |
|    | TGTAGCTAGT  | CCATATACAC | CAACGCCCTC | TTTTAATATA | TTGGAATAAA  | ACAATGTTAA | 720  |
| 25 | TACAACACGC  | CAAAGGCCAG | TATTAATCAA | TATGCAGGCA | AATTGGATGA  | TTATAATAAA | 780  |
|    | TGGAATTTCT  | TTAGATTGTT | TCAAGAATTC | CCAAGTTTCA | GAAAAATCTT  | CTTTGAGTG  | 840  |
|    | TCTATCAATC  | ATGTGTTTAT | TTGTATATTT | TAAAAGTGCA | TTAAAAATAA  | ATCCTAAAAA | 900  |
| 30 | TAGCAATATA  | CTACAAATAA | AAAAGACGCC | AACATTACCA | ACTAGTATTA  | CAATGACACC | 960  |
|    | AATTAAGACA  | GGTAAAAATA | TATTTGAGCC | TCTTTGCAAA | CTATOGATTA  | ACGCATTACT | 1020 |
| 35 | TGTTGCTAAA  | TGCTCCTCAT | CAATAATTTT | AGGAAGAATT | GCCCTAAAAG  | CAGGATCCGT | 1080 |
|    | ATAGCAGTTA  | ATAATGGTGA | TAGCTGTAGA | TATGGTTAGA | AGCGTCAGAT  | AATTTAAATT | 1140 |
|    | TGATGTTATT  | GCAAGTAAAG | GAATAATTAT | TATAATCAAA | CTTAGTATAA  | GATCAGATAG | 1200 |
| 40 | ATAAAGTATT  | TTCTTTTAT  | TATGTTTATC | AGAATATGCG | CCACCGAAAA  | TACCAATAT  | 1260 |
|    | AATAAATGGA  | AGTGTTTGAC | TCATAACCAT | CATTGATAAT | TTTAAAGATG  | ATTGGTTTGT | 1320 |
|    | CAATTCAACA  | GTAACCAAAA | TTATTTGTAA | CGAAAAACAG | ACAAAAACAAC | TCCGACGTAA | 1380 |
| 45 | GAATTACCA   | ATCAATAAAT | ATGTAAAGTT | TCTATTTTTT | AAAACCTCTA  | AATACAACAT | 1440 |
|    | ATTTATCACC  | TCTCATAAAA | ATAATTGAAT | GCATCCACCA | GCTTTTTTAT  | ACCTTCTTCT | 1500 |
| 50 | AAACTCTCTT  | TATCCAAAGC | GCAATTAATT | CTAATATAAT | TTAGTCAGTT  | AAATATCAAT | 1560 |
|    | TATTTCCGAA  | TATACATACT | ACTTGAAACA | CCATACATAA | CCCCCAAAAT  | GACTACTCAG | 1620 |
|    | AGGTATATTT  | CTACTAATTA | TGATTATATT | AAATATGAAA | ATATTATCAA  | AAAAATCAAA | 1680 |
| 55 |             |            |            |            |             |            |      |



TCATTTCTCAT CATATTTCTAG GTTGTTTTTT ACAAACTAAA TATAGTGAAT GCAAAATCAAC 1800  
 TATTATTAA ATTATGAATT ATTTTAATTC TTTCTTCTAC GAGCCAATAA CATTAATCCA 1860  
 5 GCAATTCCAA TTATACTACT AAAGATCAA CCTTTTTCG TGTCTTCTAA ACCTGTTTTT 1920  
 GGTAATTCCTG CTCGTTTTTT CTCTTGATTA GCTACTGATT CTTTAGCAAT TTTAGATTTT 1980  
 TTAACTTTAT CATTTTATC CATGAATGA ACTGGGCCAT TTGGTTTTGC TCTGTCITTC 2040  
 10 GATAATCCTG GATTGTTAGG ATTTACTGGG CCACITGGAT GAGTGTGCTC GCTCGGCTTC 2100  
 TCTGGGTTTT CAGGTCCTTT TGGATCTTTT GGTTCCTCTC CACCGAACTC TACAATCTTA 2160  
 TCTACTGGTT GITTGTGAT CTCTTCTGTT GGTGACCTC CGCCAACTTT TTCACCTGTT 2220  
 15 AATGGGKTCA CTGTGAWTGG TGTGTGAWT GTCyTACTTC CTGGKGTGCC TTCTTGTTTC 2280  
 ACTCGCTCTT CACCAGTTG TAAT 2304

20 (2) INFORMATION FOR SEQ ID NO: 433:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 4733 base pairs  
 (B) TYPE: nucleic acid  
 25 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 433:

30 GTCTTAATAT TTTACAAGAT ACAGCCTGAC AATATTATA TTACAAACCA ATATGGTAAA 60  
 GAGGCTGTGT CTCGATTAAA AAGAATTTTA GAGCAAAGTA TACCTAATAA AGAATATATT 120  
 35 GATTTATCAG AGATATCATT AATTACATTT TTTAAACAG AATATAATAA ACAACATGTG 180  
 AATGAGGAGT TTAATCTTTT ATCTTCTTCG AAAATAAAGA ATTTAGCCAA AGAAAATGAT 240  
 GGTATTATTA TTTATTTAAA TTTTAAGGAT GTTAATTATC AACTGaTAaG TGAAGGCAAT 300  
 40 ACTTTTTTTT CTGCAAAACC ATATTGTCAT TGTGTGTTCA ATAAAGATAT TTTAAATATG 360  
 GATAAAGTTG ATATAGAAAA TTTTTTCAAA AGTATAGAAG TCAAATACAG TAAATAAGAT 420  
 CAATTATTGA ATCCACCGAC CAGCAGATTT TTTTAAAAAA TTAATACCCC GTTAATACCC 480  
 45 TTTGCTTCAA TTGATGAAA ATCAATGAAA TTCAAAATTG AAGAAATCCT TTAATATCAA 540  
 GGTGTACGAC AGTCTATATA ATCATGCGAA ATTCTAAAAA TTTCTGaCGT AAAAAACAA 600  
 50 ATTCyTAAAG CAGCTCGTCG TTCACCTCAA TTCTCAAAAC GTTAATTGTG GGACGATATA 660  
 TATACAAAAC ACCTCGATGT TATGTCGAGS TGTTTTTTTT CGTTTGTGTG GGAATATGG 720  
 AATAGAGTAT AAATGAATTA ATGTCTCAGG TATAGAATTA ATTCAACTAT GAATTATTGA 780

55

|    |             |             |            |             |             |             |      |
|----|-------------|-------------|------------|-------------|-------------|-------------|------|
|    | TTAGCTTGGC  | GAGACCGTAC  | GATAGTGTTT | TAATAATAGA  | TATTCAAGGG  | AAAGAGCTAT  | 900  |
|    | CGGATGATAA  | CCAGATGATA  | ACTCAAGAAC | CATGTTTTAA  | GAGTGTATTC  | AAAAAAGGA   | 960  |
| 5  | GTGACTTAAT  | TGAAAGATT   | TATATTTTGG | AAGAAAGCTT  | TTATTyCTGT  | TTATTTTATT  | 1020 |
|    | GTGGGATGCG  | TGGTGTTCCT  | ACTTTTTAGG | TTTATATTA   | AaACAGATAA  | CTTTTCTATA  | 1080 |
|    | TATTTAATGA  | GTATCTCTCT  | AATTTGTTTA | GGAACTGCTT  | CTATCATTTA  | TAACTATAAA  | 1140 |
| 10 | ACCAATCGAT  | AAATATAAGT  | ATGAAATGTA | TAAGTCACAG  | TTAGATCTAA  | GTCTTGCTGT  | 1200 |
|    | GGCTTTTAT   | ATCTCCATAA  | AACGACAAAT | TCAAGCCCGA  | CATAAAACAG  | CATTTTCAGC  | 1260 |
|    | CCACCATAAA  | ACGACAATTT  | CAGCCCCGCA | TTGACTTAAAC | ACCACATCCC  | AAAAATATCG  | 1320 |
| 15 | TAACAATCCT  | CTACATCAAT  | CAATCCAACA | TCCCTCATAA  | TCACAACGCA  | CAAAATCTAT  | 1380 |
|    | TCATGCATTT  | TTGGAATACT  | TAGTATTTGA | AATAACGATT  | TTTATTTCATC | TAACAAGGT   | 1440 |
| 20 | TATATAATGT  | ACTGAAGGCA  | ATTTTATGT  | ATTACAAATC  | TAATCGTACA  | TGTAAAATTT  | 1500 |
|    | TGATAAACAT  | CATTAAITTT  | CGCTAACTAT | CATGAGATTA  | CAAAACACAA  | AGTAATTACA  | 1560 |
|    | TGTTAAACAC  | ATCTATACAT  | TGCTTTTGA  | GGGAATGTaa  | TATAAATGAC  | TGATAAAAAAG | 1620 |
| 25 | TACACTGCAG  | CCGATATGGT  | TATTGATACT | TTGAAAAATA  | ATGGGGTAGA  | ATATGTTTTT  | 1680 |
|    | GGTATTCCGG  | GTGCAAAAGT  | TGACTATCTA | TTTAATGCTT  | TAATTGATGA  | TGGTCTGTA   | 1740 |
|    | CTTATTGTCA  | CTCGTCATGA  | ACAAAACGCG | GCAATGATGG  | CGCAAGGTAT  | TGGAAGATTA  | 1800 |
| 30 | ACAGGCAAA   | CGGGTGTAGT  | ACTTGTTACA | AGTGTCTCTG  | GTGTAAGTAA  | TTTAACTACT  | 1860 |
|    | GGATTATTAA  | CCGCAACATC  | TGAAGGGGAT | CCTGTATTAG  | CGTTAGGTGG  | CCAAGTGAA   | 1920 |
| 35 | CGnAnATGAT  | TTATTACGAT  | TAACGCATCA | AAGTATTGAT  | AATGCTGCGT  | TATTTAAATA  | 1980 |
|    | TTCATCAAAA  | TATAGTGAAG  | AAGTACAAGA | TCTGTAATCA  | TTATCAGAAG  | TTATGACAAA  | 2040 |
|    | TGCAATTCGA  | ATTGCTACTT  | CAGGAAAAAA | TGGGCGAAGT  | TTTATTAGTA  | TTCCGCAAGA  | 2100 |
| 40 | CGTTATTTCT  | TCACCAGTTG  | AATCTAAAGC | TATATCACTT  | TGCCAAAAAA  | CAAAITTAGG  | 2160 |
|    | AGTACCGAGT  | GAACAAGATA  | TTAATGATGT | CATTGAAGCG  | ATTA AAAATG | CATCATTTCC  | 2220 |
|    | TGTTTTTATTA | GCTGGTATGA  | GAAGTTCAAG | TGCAGAGAA   | ACAAATGCCA  | TTCCGCAAAAT | 2280 |
| 45 | AGTTGAGCGC  | ACGAATTTTAC | CAGTTGTAGA | AACATTCCAA  | GGTGACAGGT  | TAATTAGTCG  | 2340 |
|    | TGAATTAGAA  | AATCATTTCT  | TCGGTCGTGT | GGGCTTATTC  | CGCAATCAAG  | TTGGTGATGA  | 2400 |
| 50 | ATTATTACGT  | AAAAGTGATT  | TAGTTGTTAC | AATCGGTTAT  | GATCCAATTG  | AATACGAAGC  | 2460 |
|    | TAGTAACGTG  | AATAAAGAA   | TAGAAACACA | AATTATCAAT  | ATTGACGAGT  | TCAAGCTGAA  | 2520 |
|    | ATTACTAATT  | ATATGCAACC  | GA AAAAGAG | TTGATTGGTA  | ATATTGCTAA  | AACGATTGAA  | 2580 |

## EP 0 786 519 A2

|    |  |      |
|----|--|------|
|    | CAATTAAGAA CACATATTGA TGAAGAACT GGTATTAAAG CGACGATGA AGAAGGAATT    | 2700 |
|    | CTACATCCAG TGGAAATTAT TGAATCTATG CAAAAGGTAT TAACTGATGA TACTACTGTA  | 2760 |
| 5  | ACAGTTGATG TTGGAAGTCA CTATATTTGG ATGGCACGTA ATTTTCAAGG TTACAATCCA  | 2820 |
|    | AGACATTTAT TATTTAGCAA TGGTATGCAA ACGCTTGGTG TAGCATTACC GTGGGCAATT  | 2880 |
|    | TCAGCTGCAC TTGTGCGCCC TAATACGCAA GTTGTGTCCG TTGCTGGCGA TGGTGGCTTT  | 2940 |
| 10 | TTATTTTCAT CACAAGATTT AGAAACGGCC GTACGTAAAA ATTTAAATAT CATCCAGCTT  | 3000 |
|    | ATTTGAATG ATGGAATAA TAACATGGTT GAATTCACAG AAGAAATGAA ATATAAACGT    | 3060 |
| 15 | TCGTCAAGTG TAGACTTCGG TCCTGTAGAT TTTGTAATAA ATGCAGAATC ATTTGGCGCG  | 3120 |
|    | AAAGGTTTAC GAGTTACTAA TCAAGAAGAA TTAGAAGCGG CAATTAAAGA GGGCTATGAA  | 3180 |
|    | ACAGATGGTC CAGTATTAAT TGATATACCT GTAAATTACA AAGATAATAT CAAACTTTCA  | 3240 |
| 20 | ACAAATATGT TACCTGACGT ATTTAACTAA AATAAAGATA AATGTTAAAG AGGAGTGGGA  | 3300 |
|    | GATTTTATGA CTAATGTTTT ATACCAACAT GGTACATTAG GCACATTAAT GGCAGGATTA  | 3360 |
|    | TTAAAAGGAA CTGCATCAAT AAATGAATTA TTGCAACATG GTGACTTAGG TATCGCTACA  | 3420 |
| 25 | CTAACAGGTT CAAACGGTGA GGTAAATCTT TTAGATGGA AAGCTTACCA TGCAAAATGAA  | 3480 |
|    | CATAAAGAA TTGTAGAATT AAAAGGTGAC GAGTTAACAC CATATGCAAC TGTAACATAA   | 3540 |
| 30 | TTTGTAGCAG ATACAAGCTA TGAACCGAAA GATAAATCTT CAGAAGCAGT TTTTGCAGAA  | 3600 |
|    | ATTAAGGAAA AGATGTTGAG TGAATAATTTA TTTTCAGCAG TAAAAATTTC AGGCTTATTT | 3660 |
|    | AAAAAATGTC ATGTACGTAT GATGCCGGCT CAAGAACCAC CTTATACAGC TTTAATCGAT  | 3720 |
| 35 | TCAGCTAGAA GACAACCTGA ACAAACTGAA ACGTATGTCA AAGGTTCACT CGTTGGTTTC  | 3780 |
|    | TTTACACCAG AATTATTCCA TGGTATCGGA TCAGCAGGAT TTCATGTACA CTTTGCGAAT  | 3840 |
|    | GATGATCGTA ACTTTGGTGG ACATGTCTTA GATTTTGAAG TAGAAGATGT TAAAGTAGAA  | 3900 |
| 40 | ATCCAAATA TAGAAACATT TGAACAGCAT TTTCCAATTC AAGATAAAGA TTTCACTAAA   | 3960 |
|    | GCAAAATATG ACTATAAAGA TATTGCAGAC GAAATTAGAG AAGCTGAATA ATGAATCCAG  | 4020 |
|    | AAATATAATG ACGGTTTATG AAAATTGACT TCATAATGCG CGATTTAGAA ATGATAGTTT  | 4080 |
| 45 | GTAATATGA TTAACCATGA CTACAATAGA ACAAATATAT TTATAATTAC GTCTAAGTAA   | 4140 |
|    | TAAAAATAAT CCCTTCACTA TTAGCAGTAG TAGAGGGATT TATTAGGTTT CAGATATTTG  | 4200 |
| 50 | AGATTTGCTG TTAATGTTAG ATTATAAAAT GTGGTACACA CTCATATAAA ATTTACTATT  | 4260 |
|    | GTATAGGCCA ATCTGTTACT ACGAGAAGCA AACACAATA ATTTACAAGT TCAATACTA    | 4320 |
|    | AAAGACAAA CGCCAAATTT TCAGCGCTTG CCCTATAAAA CTATTTTCAA ATTATTATT    | 4380 |

55

TTTCTTTAGA TICTACTACTT TTTTATTACC ATCAITCAA GTAAGCGTAT AAGTGTCTGT 4500  
 TTGGGCATTA TTAATTTTT CTGTGTGAAC ACCACGTTGA GAAGCTAATT CATTTTTTAC 4560  
 5 TTTACTGTCA ATTCTTGAT AAAGAACATT TTTATTTCT GGAAGATAA AGTAAGTTCG 4620  
 ATGTAATGCA GTAATACCAT CTACTGAAAT TGTGTAAGGA ACAGTGTGAT AACCATCCAC 4680  
 10 AGTCATTCTT TTATAGCCGT TATTACTATC TGCAGATGCT TCGTGACTCG GTA 4733

(2) INFORMATION FOR SEQ ID NO: 434:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1284 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 434:

AATAAATGGT AATTTTATAT CATCAACGGC TACAAATGGG GCAGTTTCAT ATATAAATAA 60  
 ACACATCTAC TCAACTGnTC TTAATTAATT AATAAATGAT TTACCAGATG ACATATAACA 120  
 25 GCCATTATTA GntTAATCT AATAGTTTAT TTAATTTTC TTCGGTTGTC GCCCAACTGG 180  
 TTGCGAATCT AACACACGA TGTGATCAT CGTATTTTC CCAACACGA AATTTAACTT 240  
 30 TTTGTTCTAA CTCTGCTATT TTCTCGTTAC TTAATAATAA AAATTGTTGA TTGGTTGGAG 300  
 AATCAAAGTA AAGACGATAG CCTTTATTTT TAAACCCGTC TTTTCATCTTA TTTGCCATTT 360  
 CGATAGCATG TCTGCTTATA TTAATAATAA AATTGTCCGT AAATAATTCT AAAAATTGTA 420  
 35 TGCCTGTGTA CGTCCTTTT GCTAAAAGGG CACCGTGGAT GCTTGATTGC AGTGGTAAAT 480  
 TGTTCGGTTT CATTATTTTT GGTAAAAACA ATGGCTTCCC GGCATAATGC ACCTATCTKc 540  
 GTACCACTTA TATAAAATAC ATCACAATAT TTAGCGrTgt CTTTAAATAGT CATATCTGAT 600  
 40 TGGTCACTCA TCAATCCATA CCCTAATCGT GCACCATCCA TAAATAATGG AAGCTGATAT 660  
 TGCTTACATA CTTTGCATAA CTCTTCCAAT TCTGatTTAG AGTATAATGT GCCATATTCT 720  
 GTAGGATGAG AAATATATAC CATTCCTGGG AATACCATAT GGTCCTTTTT AAAATCACTT 780  
 45 TTAATGTGCT CCATGTAAGT TTCAACATCT GAAGCACTAA CTTTTCCTTC CTTAGAGGGT 840  
 ATAGTAATTA CTTTATGTCC ACTATATICA ATTGCACGCG CCTCATGCAC AGCAACATGA 900  
 50 CCAGTGTCTG CTGAAATGAC CCCTTCGTAA CTTTCTAACA TTGAATTAAT AACAACTGA 960  
 TTGGTTTGTG TTCCACCTaC TAAAAACGA ATTGTAGCAT TTGgCAGTC AATTGTATCT 1020  
 TTAATCTTTT CAATTGCCTG AGCTGTGAAT TgATCAAAGC CATATCCCGA AGCTTGTA 1080

TCGAATGAAA TCACACATT TCCCCTAAA ACTAATATCA ACATTTTAAAT AAGATAAAACC 1200  
 AATTTCAAAA CIAGTTCGAT ATTTAAAAATG TATTATGGAT GGNATAAGTT TGTATCGCAT 1260  
 5 TATCGCGAAG TTGnATAAAT ATAT 1284

(2) INFORMATION FOR SEQ ID NO: 435:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1072 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 435:

ACAGCTTTTG GTAArgGAGA AAwTCAaTat GAAACAGTAT AArGCGTATT TaATCGATTT 60  
 20 AGATGGCACA ATGTATATGG GAACAGATGA GATTGATGGA GCAaAACAAT TCATCGATTA 120  
 TTTAAATGTA AAAGGCATTTC CTCATTTATA CGTAACTAAT AAITCAACAA AAACACCTGA 180  
 GCAAGTAACT GAAAAATTAC GTGAAATGCA CATTGATGCT AAACCAGAAG AGGTTGTAAAC 240  
 25 GTCAGCGTTA GCCACTGCTG ATTATATTTC AGAACAAATCA CCAGGAGCAT CAGTATATAT 300  
 GTTAGGTGGG AGTGGTTTAA ATACTGCGTT AACCGAAGCG GCACTTGTC AAAAAATGA 360  
 CGAGCATGTT GATTATGTAG TTATTGGACT TGACGAACAA GTTACATATG AAAAGCTTGC 420  
 GATTGCAACG TTAGGTGTAA GAAATGGtGC AACATTtATT TCTACAAATC CTGATGTATC 480  
 AATTCTTAAA GAGCGTGGTT TATTACCTGG TAATGGTGCT ATTACAAGTG TTGTAAGTGT 540  
 35 ATCGACAGGT GTATCGCCAC AATTtATTGG TAAACCAGAA CCGATTATTA TGGTTAAAGC 600  
 ATTAGAAATT TTAGGATTAG ATAAATCCGA AGTTGCTATG GTAGGCGATT TGtACGATAC 660  
 CGATATTATG TCTGGTATTA ACGTAGGTAT GGATACGATT CAiGTACAAA CAGGTGTATC 720  
 40 TACGTTAGAA GATGTGCAAA ATAAAAATGT GCCACCAACG TATTCTTTTA AAGATTtAAA 780  
 TGAAGCAATA GCTGAATTAG AAAAAATAGAT ATAGTCATTt TATAAGTAG GTGAATTGAT 840  
 TTGGTAAAAA TAGTTGTTTC GAGGAAAAAT CCAGATAAAT TTTATCAACA ATtAAGTAAA 900  
 45 CTtGGTGACG TTGTTATGTG GCAAAAAATCA TTAGTGCCTA TGCCTAAAGA TCAATTtGTG 960  
 ACaGcCTTCG TGACGCAGAT GCTTGTtTTTA TTACATTAAg TGAACAGATC GATGCAGAAA 1020  
 50 TTTTAGCGCA ATCACCAAAT TtAAAaGTAA TTGCGAATAT GGCTGTAGGA TA 1072

(2) INFORMATION FOR SEQ ID NO: 436:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 3271 base pairs

(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 436:

|    |            |            |             |             |            |             |      |
|----|------------|------------|-------------|-------------|------------|-------------|------|
|    | TAAAAACTTA | CTTTAACACC | ATTCCTTTTT  | AACCTTTTTTC | GTGTTTCnTT | TnCTTAAGTC  | 60   |
| 10 | CATCCATATT | TTTAATGATG | TCATCTGCTG  | TTTATCTTTT  | TAAATCTAAC | ACTGaGTGAT  | 120  |
|    | AACGAATTGG | TAGCACAGGA | TCAAATCCTT  | TATGGAATCC  | AGTATGTTCa | AATCCTAAGT  | 180  |
|    | TACTCATTTT | ATCAAGAAGC | CAATCATTAC  | CAGCATTACC  | TGTAATCTCG | CCATCATGAT  | 240  |
| 15 | TCAAGTATTG | ATATGGTAAA | TATGGATCGA  | TATGTAGGTA  | TAGACAACGA | TGTTTTTTAA  | 300  |
|    | CATATTTTGA | TAATTCATTA | AAGAAAAAGT  | GTACGAGTTC  | TTGATTTTCA | TAATCAATCA  | 360  |
|    | CTGGACCGCG | ATTTGAATAA | AAATACTTGA  | ACACTTTCAT  | AACAGGTACA | CGAGTAAGTA  | 420  |
| 20 | AGCAAGCTCG | AATGACCTCG | TTATTATTGT  | TTTTTATTCC  | CACFAAATGT | GTTCATAAAC  | 480  |
|    | CTTCAGCAAG | CTTTAACTCA | TAGTGGCCAA  | CAGTTTGCGT  | GaaATGACTG | TATGGCATGC  | 540  |
|    | TATCTGTAAA | GGCACCAAAC | TCTTTAGCTG  | TTAAATTTGT  | AAACTTCATT | ATCATTACTC  | 600  |
| 25 | CTATTGTCTC | CTCGTTAATT | AATTTCATTT  | CCGTATTTGc  | AGTTTTTCTA | TTCCCCCTCT  | 660  |
|    | GCAAATGgCA | AAAATAATAA | ATCTAATCTA  | AATAAGTATA  | CAATAGTTAA | TGTTAAAACT  | 720  |
| 30 | AAAACATAAA | CGCTTTAATT | GGGTATACTT  | TTATAGTAAT  | ATTAGATTIT | TTGAATACAA  | 780  |
|    | TTTCAAAAAA | AGTAATATGA | ACGTTTGGGT  | TTGCTCATAT  | TACTTTTTTT | GAAATTGTAT  | 840  |
|    | TCAATTTTAT | AATTCACCGT | TTTTCACITT  | TTCAAAACAGT | ATTCGCCTAA | TTTTTTTAAA  | 900  |
| 35 | TCAAGTAAAC | TTAATTATTG | AATGTTTGTT  | GGATAGATTG  | TAAATATTTA | ATGATTTCCT  | 960  |
|    | CACGCGTGTT | AGATTTAAAT | CGCTTAAACGA | TTTGGCTACC  | AATGACAATG | CCATCTGCAA  | 1020 |
|    | CCTCTTTTAT | ATCTGCAACA | TGTTGTGGTG  | TTCTTATACC  | AAATCTGGCG | ACAATGGCA   | 1080 |
| 40 | CATTGGCTAT | CGCTTTAATT | GACTCAATTT  | TTCGTTTTAA  | TTCTGGATGA | AACGCACCGT  | 1140 |
|    | TTTGCCCTGT | TGTGCAATTC | ATCGTCACAG  | TATAAATAAA  | GCCTTCCGCA | TGGGATACGA  | 1200 |
| 45 | TATCTTTTAT | ACGTTTGTCa | TCAGTAGTCA  | TGCAACTTAA  | CGATATGATT | TTGACGCCAT  | 1260 |
|    | AGTGACTAAA | TTGTTGTTTT | AAACGCTGCG  | ATAATTCTATA | TGGTAAATCA | GGAAATAATTA | 1320 |
|    | AGCCGTAGAC | ACCAGTATCT | CGACATTTTT  | CAAAAAACGC  | TTGTTCTCCA | TAATGACAAA  | 1380 |
| 50 | TAATATTATA | ATAGTTCATT | AATACATAGT  | TACACTTAAT  | TTGATCACCA | TGTTTTTCTA  | 1440 |
|    | ATTGATIGAA | AATATAATCT | ATCGTGATGC  | CTTGTTAAAT  | CGCTTGTTGA | CCTGCTTCCA  | 1500 |
| 55 | TGATAACTGG | ACCATCAGCA | ACCGGATCAG  | AGAAAGGTAC  | TCCAATTTCA | ATTATATCTG  | 1560 |

EP 0 786 519 A2

|    |   |      |
|----|---|------|
|    | GTATAAATAA TTAGTTCATT TGCAAGACCT CGCTCTACCA TATATTGTCT AATTGTTTC  | 1680 |
|    | ATATCTTTAT CGCCACGTCC AGAAATAGTT ACTACAATAA TATCTTCTTT CGACATCGTA | 1740 |
| 5  | GGCGCTAGTC TTTCACATA ACTCAGTGCA TGTGCATTT CAATTGCAGG TATAATACCT   | 1800 |
|    | TCATGTTTG TAAAGTTGAT TAAAGCATTG ATTGCTTGTC TATCACTAGC ATTTTCAAAA  | 1860 |
| 10 | GTTACTCTAC CAATGTCGTG GTAATAAGAA TGTTCTGGTC CAATACCAGG ATAATCAAGT | 1920 |
|    | CCTGCTGAAA TAGAATGTGC TAGTTGCACT TGCCCATCTT CATCTTGAAT TAAATACATT | 1980 |
|    | TTAGTACCAT GTAATACGCC AGGTGATCCT TTGCCAATTG CAAGTGCATG TTTATCAGTA | 2040 |
| 15 | TCATCGCCTT GACCTGCGGC TTCAACACCG TATAATGCAA CATCATCTTT AATAAATGGA | 2100 |
|    | TAAATGTAC CGATTGCATT TGAGCCACCA CCGATACATG CTACAATTGC ATCCGGAAGT  | 2160 |
|    | CGACCTTCTT TCTTCAATAT CTGTGATTTT ATTTCTTTAC CAATCACACT CTGAAAATCT | 2220 |
| 20 | CTAACAATCG TTGGGAACGG GTCTGGACCT AATGCAGAAC CTAATAAATA ATGTGTATCA | 2280 |
|    | TCTACATGAC TTACCCAATA TTGCAATGCT TTATTAACTG CATCCGATAA AGTCCCTTGA | 2340 |
|    | CCATCTTCAA TGCCACAAC CTTTGACCA AGTAATTCCA TTCTAAATAC ATTAAGTTGT   | 2400 |
| 25 | TGCTTTTTAA TATCTTCACT TCCCATAAAG ACAACAAGTT CCATATCAAA TAATGAGCA  | 2460 |
|    | ACCGTAGCAC TAGCTACACC ATGTTGACCC GCACCAAGTT CAGCAACAAG CTTCTCTTTC | 2520 |
| 30 | CCCATCTTCT TAGCAAGCAA CGCTTGACCT AACGCATTAT TAATTTTATG GGCCTCTGTA | 2580 |
|    | TGATTTAGAT CCTCTCGTTT CAAATATATT TTAGCGCCAC CTAGGCTTTC AGTATATGAT | 2640 |
|    | GCAGCATATG TAAGTGTGTG CGCGCGTCTC ACATACTCTG ATAAATAGTA TTCCAGTTCT | 2700 |
| 35 | CTTTGAAACT CTGGGCTCTG TTTTGCTCTT TTATAAGCTT TTTTCAACTC AATAATGTCT | 2760 |
|    | GGCATTAATG TTTCTGGAAC ATATTGCCCT CCATATTCAC CAAAGAAACC TAATTCATCT | 2820 |
|    | GCTTCGTGTT GATTGTTT ATTCAATTGC TCTATCTCCT TTCACAATAT TTACAATTGC   | 2880 |
| 40 | TGTCATTTT TCTATATCIT TTGCCCCATT TACTTCTATA CCTGATGCAA GATCATAACC  | 2940 |
|    | TTGATGTGAT AATTAAAGTT GATTAACTGT TTGAATATT TCAGAGTTAA TGCCTCTGTC  | 3000 |
|    | TATCAAATAA GGTATGTCTT TTATGTGCTT CAAAATAGTC CAGTCATATG TTTGACCGGT | 3060 |
| 45 | ACCACCATAC GACACTGAGG GTGTGTCGAT AATAAATAA TCTACGAACC CTTTATATTT  | 3120 |
|    | ATTATGTTT TGGATTATGT TTTTCATCTG CAGTAAAGCT TTAGTGATTT TAATGCTTGA  | 3180 |
| 50 | ATATKCTTTT TTAATTTTCT GTATAAAATC AATGATTCT GtGTAACGT ATTGTGkTAA   | 3240 |
|    | WTGaCGwATG CtTAawACgT GTGCCAATGG T                                | 3271 |

(2) INFORMATION FOR SEQ ID NO: 437:

55

## EP 0 786 519 A2

(A) LENGTH: 1553 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 437:

|    |   |      |
|----|---|------|
| 10 | ATAATGAAAT AATACTGTGT TTTATCTGCG AAATGTATCA TTTCTAATC GTTTCACAGT  | 60   |
|    | AAAATGAAA GATAAGTGT GTTTTACTT GAATTTTGAC TAAAATTACT CTATATTAT     | 120  |
|    | TAATTGAGCT ATGCTTATTA TTACAATTG ATTACAAAT TTAATTTGT TAATTGAATG    | 180  |
| 15 | ATAATATTAA ATAAGAAAC TTACACAAGC AATATGAGT TGAGCCCAA AATACITGTT    | 240  |
|    | AAATCAAAGT TGAAGCTAC AAATAATGAA AATTATAAAC TTGAATCTGA AAGTAATTAC  | 300  |
|    | TATAATTATG ACAATGTTAA CTTTAAACG CACTTATTAA TTAACACAT AATGTTAATA   | 360  |
| 20 | TCTAATTTAT TCAAGTACTT TCGCAAGATT TATTATCTAA ATAACGGGG AAAGAATCAT  | 420  |
|    | GAGTtCACAA AAAAGAAAA TAGTCTTTT GCGTCTTCT TATTAACCGT AATAACGATT    | 480  |
|    | ACCTTGAAGA CGTATTTTC TTATTATGTT GATTTTCTT TAGGTGTTAA AGGTTTAGTA   | 540  |
| 25 | CAAAACTTAA TATTATTGAT GAATCCATT AGTTTAGTAG CACTGGTTT AAGTGTGTTT   | 600  |
|    | CTATTCTTTA AAGGCAAAA AGCATTTTGG TTCATGTTCA TAGGCGGCTT CTTATTGACG  | 660  |
|    | TTCTATTAT ATGCCAATGT TGTTACTTT AGATTCTTCT CTGATTTTTT AACGTTTAGT   | 720  |
|    | ACTTTAAACC AAGTAGGTAA CGTGAATCT ATGGTGGTG CGGTTAGTGC ATCATTCAAA   | 780  |
|    | TGGTATGACT TTGTTTATTT CATTGATACG TTAGTTTACT TATTCATTTT AATATTTAAA | 840  |
| 35 | ACAAAATGGT TAGACACAAA AGCATTTAGT AAGAAATTG TTCTGTCTGT AATGGCGCT   | 900  |
|    | TCAGTAGCAT TATTCTTCTT AAACCTAGCT TTTGCTGAAA CTGACAGACC AGAATTATTA | 960  |
|    | ACACGTACAT TTGACCATAA ATATTTAGTG AAATATTTAG GACCTTATAA CTTTACAGTA | 1020 |
| 40 | TACGaTGGTG TTAAACTAT CGAAAAAAT CAACAAAAAG CGTAGCATC TGAAGATGAC    | 1080 |
|    | TTAAcAAAG TATTAAATTA TACGAAACAA CGTCAAAChG AGCCTAACCC rGAWTATTAT  | 1140 |
|    | GGGGTGcAA rGAAGAAAA TATTATTArG ATTCAATTAG AAAGTTTCCA AACCTCTTTA   | 1200 |
|    | ATTAATAAAA AGGTTAATGG TAAAGAAgTA ACACCGTTTT TAAACAAAT ATCAAGTGGG  | 1260 |
|    | AAAGAGCAAT TCACATACTT CCCTAACTTT TtCCATCAAA CAGGTCAAGG TAAACATCT  | 1320 |
| 50 | GACTCTGAAT TTACAATGGA TAACAGTTTA TACGGTTTAC CGCAAGgTTC TGcCTTTTCA | 1380 |
|    | TtaaaaGGAG ATAATACGTA TCAGTCATTA CCAGCAATTT TAGATCAAAA GCAAGGCTAC | 1440 |
|    | AAATCTGATG TCATGCACGG TGACTATAAA ACATTCTGGA ACAGAGACCA AGTATATAAA | 1500 |

55



(2) INFORMATION FOR SEQ ID NO: 438:

## 5

- (A) LENGTH: 1419 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

## 10

|             |             |             |             |             |             |      |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| ACAAAATCAA  | TCAGCAAAAT  | GATTGGATTA  | AGACGCAAAAT | TGAGCGTTCA  | ATGGAAGCGC  | 60   |
| AAACAGTTGG  | CATTAAATGAT | CAAAATATAG  | AAATATATAG  | TGAAcGtGg   | GATTTATACC  | 120  |
| ATACACTCGT  | ACCTTTAAAT  | CAGAATATGC  | ATAAGTTGGC  | ACTTAAAACT  | CAAAACCTTA  | 180  |
| CCAATGAAAA  | TTATAATATT  | AATGATGTGA  | AAGTTAAAAA  | GATTATTGAA  | GATGAACGTC  | 240  |
| AAAGACTAGC  | ACGAGAAGCTT | CACGATTCTG  | TTAGTCACGA  | ACTTTTTCG   | GcmAGTATGA  | 300  |
| TGCTATCTCG  | TATCAAAAGAA | ACGAATTAGA  | ACCACCATT   | GACCAACAAA  | TTCTATTTTT  | 360  |
| AGAGAAAAATG | GTTCAAGATT  | CGCAGTTAGA  | AATGCGTGCT  | TTGCTGTTAC  | ATTTAAGACC  | 420  |
| CTTTGGTTTA  | AGAGACAAAT  | CTTTAGATGA  | GGGTATTAAA  | GATTTAGTTA  | TTGATTTTACA | 480  |
| AAAAAAAAGT  | CCAATGAAAG  | TTGTGCATGA  | AATACAAATG  | TTTAAAGTGC  | CTAAAGGTAT  | 540  |
| TGAAGATCAT  | TTGTTCAGAA  | TTACACAGGA  | AGCAATTTGC  | AATACATTCG  | CTAATTCAAA  | 600  |
| CGGTACAAAA  | GTGACAGTAG  | AATGTTTTAA  | TAAAGACGAT  | TATTTATTGT  | TGAGAATTCA  | 660  |
| AGATAATGTT  | AAAGGTTTTA  | ATGTTGATGA  | AAAATTAGAA  | CAAAGTTATG  | GACTTAAAAA  | 720  |
| TATGCGTGAA  | AGAGCTTTGG  | AAATTTGGTC  | AACGTTCCAT  | ATTGTATCAT  | TGCCAGATTC  | 780  |
| AGGTACACGT  | ATCGAGGTGA  | AAGCACCTTT  | AAATAAGGAG  | GATTCGTATG  | ACGATTAAAG  | 840  |
| TATTTGTTGT  | GGATGATCAT  | GAATGGTAC   | GTATAGGAAT  | TTCAAGTTAT  | CTATCAACGC  | 900  |
| AAAGTGATAT  | TGAAGTAGTT  | GGTGAAGCGC  | CTTCTGGTAA  | AGAAGCAATT  | GCCAAAGCCC  | 960  |
| ATGAGTTGAA  | GCCAGATTTA  | ATTTTAATGG  | ATTTACTTAT  | GGATGACATG  | GATGGGTGAG  | 1020 |
| AAGCGACGAC  | TCAGATTAAA  | AAAGATTTAC  | CGCAAAATTAA | AGTATTAAATG | TTAACTAGTT  | 1080 |
| TTATTGAAGA  | TAAAGAGGTA  | TATCGTGCAT  | TAGATGCAGG  | TGTCGATAGT  | TACATTTTAA  | 1140 |
| AAACAACAAG  | TGCAAAAGAT  | ATCGCGCATG  | CAGTTCGTAA  | AacTTCTAGA  | GGAAGATCTG  | 1200 |
| TTTTTGAACC  | GGAGGTTTTA  | GTA AAAATG  | GTAACCGTAT  | GA AAAAGCCG | GCAAGGCTTAT | 1260 |
| TTTAAATGCT  | TACAGTAACGA | TGAAATGAAA  | TATTTATTAT  | GATTCGGAAG  | GGTTACTCAA  | 1320 |
| ATCAAGAAAT  | TGCTAGTGCA  | TGCATATATTA | CTATTAAAAA  | GGTTAGACGA  | CATGTGAGTA  | 1380 |

## (2) INFORMATION FOR SEQ ID NO: 439:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 608 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 439:

ATGTGGCGG ACGTAAAGGA CATGTTTATA CTGATGATCG AGCATTAGAT ATTGATATCG 60  
 TTCCGCTCTG TCAAGCAGAT GGTAAGCTA CTAACCCGA ACAATTATTT GCAGCAGGTT 120  
 ATGCATCTTG CTTCACGGT GCTTTCGACC TAATTTTAAA GCAAAACAAA GTGCGTGATG 180  
 CTCATCCAGA AGTAACACTA ACAGTGAGAC TAGAAGATGA TTCAGACTCA GAAAGTCCTA 240  
 AATTAAAGTG TTCAATTGAT GCGACAATTA AAAATGTTAT ATCTCAAGAA GAAGCTGAAA 300  
 AATATTTACA AATGGCTCAT GAAFTTTGTC CATATTCAA AGCGACTCAA GGAATATTA 360  
 ATGTCGATTT AAATGTAAT GTTGATGATT AGCATTAACT TAAAGAGATT ATTCAACGTT 420  
 ATTAATAAAA TTCACATAAA ATTCAAATTC stCfAcCAAA AATTTTGGT TGGYATTTTT 480  
 TTCTATTCTG GATTGAAATT TCTGGCAATA TTAACTGAA AATGATTGTA CCTTAGTCAT 540  
 CATAAATGTG ACCGGTTCa ATACTGGCTT GACTTCTTCG CATACCGTCT ACAAATAAAA 600  
 GTCCAGTG 608

## (2) INFORMATION FOR SEQ ID NO: 440:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 602 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 440:

CCTTTCAATt TCTCCCAAgA TTtACGCATG TCTGACCAAA TGtCAATACA ATACCTGAAT 60  
 AATGATTtCT GGAGTAATG CGAAAAATAT TTAAAGTTG CTATAGATCA ATTTTCAAAT 120  
 TATAGTATCT CATCTCAAGT TTCTAACTAT CATTTCACAG TATTACTtGG AGATGCCCAA 180  
 AAACCACTTA TGtATCTPAA TAAAAATCGC GGTGGTGATG GTGGCATACC AGGTTATATT 240  
 ATGATTtTATT TAGTGCCGAG TACAAGTACA ATTAATTCTA TGAAAAGCTT AATTGCACAT 300  
 GAAGTAAATC ATAATATGCG CTATCAATAT ATTGATTGGG ATGGCGGAAG TTTGATTGAA 360



|    |  |      |
|----|--|------|
|    | AAATTCITTA TAGACTGATG CAAATCTAAC ATATGAAACT TGATCAACAT GCATTAACAA  | 1200 |
|    | GTCATAACG TGTTACCTA TATCTCGTGA AGACACTTCC GTATGACCTT CATCTCGTAA    | 1260 |
| 5  | TTGCCATTCA ACCTTGTTAG TTATGTCCTC AAGTTGTIGA TATCTAACTG GTCGTTCTC   | 1320 |
|    | ACAAGAAGCG ACAAGTCCAT TAAGTATCTT TTCTCTTGAA AACTGCTCTC TTGTGCCATC  | 1380 |
|    | TTTTTTCACA ACTATAAGCT GACTAACTTC GATATGTTCA AATGTAGTGA AACGTGTTCC  | 1440 |
| 10 | ACAATTTTCA CATCTCTTC GTCTTCGAAT GGCATTAAAT TCATCGGCAT GCCTTGAATC   | 1500 |
|    | TACAACITTA GATTGTGTAG AATTACATTT CGGGCAATTC ATTACATCAC CCTCTTATT   | 1560 |
| 15 | TTGATTATGC CTAATTATAC TATAAATCTA GAGATGAAAA AAGAATCCCT CAATTTAATT  | 1620 |
|    | CATTAAACCA AATAATGAAA CAATAAAAAA CATTATATCG TTACTTATTA AGTAATTTCG  | 1680 |
|    | ATGACAATAT TATTGTATTA AAAATAAAAA ACCTAACTCC GAAGTCAGAG TTAGGCTATA  | 1740 |
| 20 | AATTAAITGT ATTAACITTC ACTTACAGTT TCTTTTGATG TCAAAAGTGC TCCAATTTCG  | 1800 |
|    | TCAGCAACAT CTACAACCTT ATTTGAATAA CCCCATTCTAT TATCATACCA AGCAATAACT | 1860 |
|    | TTTACTTTAT TCCCTGACAT GACCATTGTT GATTTTGATC CAATAATAGC TGAATTGGGA  | 1920 |
| 25 | TTAGTATTAA AATCAACAGA CACTAGTGGT TGATGTTTGA CTCTCTATGAT ACCTTCTAAA | 1980 |
|    | CCTGCATTTT CAAAAGCTTG GTTTACTTCT TCTGCAGTTA CTCTCTTTTC TAAATCAACA  | 2040 |
|    | ACTAAATCAA CGAGCGATAC ATTCTTTGTT GGTACACGTA ATGCCATGCC GTGTAATTTA  | 2100 |
| 30 | CCTCTTAATT CTGGTAATAC TTCTTTTAAA GCTTTGCGCG CACCAGTAGA AGTAGGAATA  | 2160 |
|    | ATGCTTTTAT TACATGAACG TGCACGTCTT AAATCTTTAT GTGGATTATC AATATTTTTC  | 2220 |
|    | TGGTCATTTC TAATAGCGTG AACAGTAGTC ATTAACCATT TAACATTCC AACTGATTA    | 2280 |
| 35 | TTTAAACTTT TTGCAACTGG ACCAATGCAA TTAGTAGTAC ATGAAGCATT ACTAAAAATG  | 2340 |
|    | TCAAATGCTT CTATATCTAA TTGGTTATCA TTTACGCCCT TAACACCAT TTGAACATGT   | 2400 |
| 40 | CCACCTTTTG TAGGACCAAGT TAACAAAAAT TTTTGGCAC CTGCTTTAAT ATGTGCGATG  | 2460 |
|    | GCTTTATCAC CATGATTAAA TTTACAGTTT GCATCTATAG CAATATCGAT ATCTAATCTT  | 2520 |
|    | TTCCATGGCA AGTTTTCAGG ATTGCGATCA GCAACCAATT TAATTTTATG ATCT        | 2574 |

(2) INFORMATION FOR SEQ ID NO: 442:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 3326 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear



ATCAAGGAGC GAAACATCCG AATGAAAAAG TTGTTGCTAC TGACTCAGCA ATGATTCCTT 1860  
 ATGCTGCTTG GCAGAAATTT GCTAGACCAC GCTTTGaTAA TAIGATTAGT TTTAATAGTA 1920  
 5 CCAACGAITTT ATTAACATGG TTACAGATC CATTCaTCAA AGATATGCCA GGAAACCGG 1980  
 TTAACATTAA TGATGGTGTG CCCAGGTTAG ATACTTTAAT AGACAGCCAT GTAGGTTATA 2040  
 AAAGGAAGTT AAATAGAAAA GATAACACAT ACGATACTGT ACCCTAATC AAAATAAAGT 2100  
 10 CGGTAAAGA TACAGAAATT AAAAATGGAA AAAAAGTAAA AAAGACTATT AACATAACAT 2160  
 TAGATATGGA TGGGCGAATT CCAATAAATG TTGGACAGG AGATTGATT GCACGTTCTG 2220  
 GAAGAGGAAC TTTAATTAAT CTTAATTTAG AAAATCTTGA TCGGTTGAGT AAACGTGATTA 2280  
 15 CTGGTGAAC AAGTGGTATG TTAGCAGAAT GCGTAATCTT TTTAAATGAA AGTTTTAACA 2340  
 TCTCAGAAAA TGAAAAATAA AATTTTGCAG ATAGAAAGaa ACAATTATCA GAAGGATTTA 2400  
 20 AGGATAGAT TAACTTATTT CAGTTAGAAG AAATGGAAG AACTTTAATT AGTAAATATA 2460  
 ACTCACTTGA AGAAGTTGCA GATGAACAA TAGAAAGTAT TAGTGCTGTT AAACACTTAT 2520  
 TACCTGATTT TGCATTGGAT GCATTAAAAA AAAGAATTAA TGAGTTGTTT AAAGGTATAA 2580  
 25 AATCTTTTAT AGAAAAAGTG TATGATAGTA TAGATAATGA AATTTTAGAA ATTTTCAAAA 2640  
 ATATAGATCA CGACTTCAGA GATGGAGTAT CTGAAGAAAT GATGAAACAT TTGAAAGTAG 2700  
 TGAACAGAA TATAGACCAA ATAAAAATC AAAATGATAT TTATGGTAGG CAAATTGCAG 2760  
 30 ATATTAGAAG TATTATGAAA CAACAAGATG CAACAATTTT AGATGGAAT TTTCAAAATA 2820  
 ATTGTAGCGG CGAAAAATAG GTACAGGGTC TAGTTATACC TTCTAATTAT TTAGGAAGAA 2880  
 35 AAATGAAAT ATTAAGAAGC CATATCGATG ATGGTATTAA AAAAATAGCA GACTATGTTT 2940  
 AAGGTATATA TGATGAATAT GCATCGAAAA TTGTCGATGT AATAAAATAT TTGATTATAA 3000  
 CAATTCCTAA AATACGTAAG AATTTAAGAC ATGCAATTGA AATGTTAAAT GTAAAAAAGA 3060  
 40 AAGAAATTTT GTCCCTGATT CCTAATGTAA CTTGTAATTA TATTAATACT AAATTAGAAG 3120  
 AATTAGATAA TACTTTTAGGC AAATGGGAGC CTTTCTTAA TGATTTAAAA GCAGTGCAC 3180  
 CAATTTTAGA TAACCATTTA GATGATATTG TTAAGAACAT GAAGCCTTTG ATTGTACAAA 3240  
 45 TGAYATWTGA ACCATCACAT TATGACGATA TGTTTAATTC aGAAAAAGCT TTAACGcCAG 3300  
 TGTTCCTAAG CGTTTATATA AGGTTG 3326

(2) INFORMATION FOR SEQ ID NO: 443:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 5301 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double

## EP 0 786 519 A2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 443:

|    |                    |            |            |            |            |            |      |
|----|--------------------|------------|------------|------------|------------|------------|------|
| 5  | ACCTATAGCT GGCCGTC | CA         | ATCCTGAATC | ATATGTCGCT | ATAAACCCAG | TTGCATGCC  | 60   |
|    | AGCAATCTCG         | GTTTCAATT  | CAAAATCATG | CTCTTCAAA  | CGATCTATTA | AAGTTCGAGA | 120  |
|    | CGAAATATT          | TCITCATTAC | CAAGTTCAGG | ACGTCATGAC | ATTCTATGAC | TGATTTCGAT | 180  |
| 10 | ATAACTATAT         | TTATTGTGTT | CTATATAATC | GAGAATTGTT | TGTTTTTCAC | TCATTTTTTA | 240  |
|    | CTATCTCCCT         | TTACCCCTAC | ACTCATTTTA | ATCATCCACA | ATATTTTGTT | CTTTCAAAAT | 300  |
|    | GAATTATTAC         | TTATCTATC  | GGTTTTTATC | CATGATGTCA | TCTAGTTTTT | CTTTATTTAC | 360  |
| 15 | AAAATTTTCT         | AATAACTAAA | AGCCTTTCAT | AAATTTATAA | AACAGTTTCA | AATTGTAAAA | 420  |
|    | CATAAGCCCT         | ATTGTTACAA | TTTAAGTATT | GTTAGAAACA | CAACTTACAA | CAAACTAAG  | 480  |
| 20 | TTTTATAATG         | CAACAAATCA | TAAGCGTTTT | ATAGTTTGA  | GTAGTTAATA | GGAGGAAAT  | 540  |
|    | CAAAATGACA         | AAAATGAATG | TTGAAAGTTT | CAATTAGAT  | CACTACTAAG | TGGTTGCCCC | 600  |
|    | ATTIATTAGA         | TTAGCGGAA  | CGATGGAAGG | ATTAAACGGA | GATGTCATT  | ACAAATACGA | 660  |
| 25 | CATTCGTTTC         | AAACAACCAA | ACAAAGAACA | TATGGATATG | CCCGGACTAC | ATTCATTAGA | 720  |
|    | ACATTTAATG         | GCTGAAAATA | TTAGAAATCA | TAGTGACAAA | GTTGTTGATT | TAAGTCCTAT | 780  |
|    | GGGTGCCAA          | ACTGTTTCT  | ATGTATCATT | TATTAATCAT | GATAATTATG | ATGATGTATT | 840  |
| 30 | AAATATTGTT         | GAAGCAACTT | TAAATGATGT | GCTAAATGCT | ACTGAAGTGC | CTGCTTGTA  | 900  |
|    | TGAAGTACAA         | TGTGGCTGGG | CAGCAAGTCA | TTCATTAGAA | GGTGCTAAAA | CTATCGCTCA | 960  |
|    | AGCATTCTTA         | GACAAACGAA | ACGAATGGCA | TGATGTTTT  | GGTACAGGAA | AATAAATCTT | 1020 |
| 35 | AGTCAATCAA         | GTTAATCAGA | AAAGCAGTCG | AACAATGATT | TTACAATCGC | CATTGTCCAA | 1080 |
|    | CTGCTTTTTA         | TTATGCTTCA | AAGTCAAAAA | ATCGAACAAA | TGAAAAAGTA | AAATCTTTAA | 1140 |
| 40 | CATTGTGCGG         | ATTIATTTGA | GAACCACTAT | AATTTCTTAA | TTAGTCCCAT | TAACACGAAC | 1200 |
|    | TGCATAGGTA         | ACCTTAAATA | TAGTTGCCAT | GTTGGCAATT | GTTTATCTCC | TAAAGGTAAC | 1260 |
|    | TTTTTAACTG         | CCATATAGAT | ATTAGCTGGG | AATACAGCTA | GCAAGAATAG | ATTGATTGTA | 1320 |
| 45 | TTTTTCAAGC         | ATTGAGATGG | TCTTTTAATT | AAAAGTGCAA | GTCCAAATAA | TATCTCAAAG | 1380 |
|    | ATTCCTGTAA         | CAAGAACCGC | TGTTTTTGA  | AGTGGCAAA  | ATTCGGTAT  | GATATTTCTA | 1440 |
|    | AATTGTCGTT         | CTCGTGTAAA | ATGCAATACA | CCTATTACAC | TAAACCTAT  | TCCTAATAAA | 1500 |
| 50 | TATCTTAGTA         | TGTTCAATCA | GCTTCAACTC | CTATTCTGTA | ATGATTTTAT | GAATTAAATG | 1560 |
|    | AGGCGATACA         | ACATGATCAG | CAATTGTTAT | GCTTGAATCT | AATTTTTTAA | CAACATCGTC | 1620 |

55

EP 0 786 519 A2

|    |             |             |            |             |            |            |      |
|----|-------------|-------------|------------|-------------|------------|------------|------|
|    | AATTTTTTTA  | TTTAAACAA   | TACCAACCGC | TAAATCAATA  | TCATCCTCTT | TIGTTAAACG | 1740 |
|    | TCCCGCTCCT  | AACATCATCG  | AAGCGACACC | TATATCGTTA  | GAGACTAATT | CAGTCACATA | 1800 |
| 5  | ACCTGATTTT  | TTAGCTTTAT  | ATTCAATTTG | ATATTGAGCT  | TGTGGCAAAC | GCTCTGGATG | 1860 |
|    | GTCAATAACA  | GTTTCGTGCG  | CACCTTGGTT | TTTAATAAAT  | GTTTGAATT  | TTTCTAATGC | 1920 |
|    | TGCACCTGAA  | TTAATTGCCT  | CAATTAGCAA | CGCTCTCGTT  | TCTTCAAGCG | TTTCAGCTTT | 1980 |
| 10 | GTTTGCAAGT  | ACAACCATTT  | GAGAACCTAA | TGTTAATACA  | AGTTCTGTTA | AATCTTTCGG | 2040 |
|    | ACCTTGTCTT  | TTCAACGTAT  | CAATTGCTTC | TTGTAACCTA  | AGCGCATTGC | CAATCGCACG | 2100 |
| 15 | TCCAAGTGGC  | TGATTTCATAT | CAGAAATAAT | CGCCATCGTA  | TTACGTCCCA | CATTATTACC | 2160 |
|    | AATACGTACC  | ATTGCGTGGC  | CTAATGCTTC | AGCATCTTCT  | AATGTTTTCA | TAAATGCACC | 2220 |
|    | GCTACCAAGT  | TTTACATCTA  | ATACAATTGC | ATCTGCACCA  | GCAGCAATCT | TTTACTCAT  | 2280 |
| 20 | AATTGAAGAG  | GCAATTAATG  | GTAATTGAAT | GACAGTACCA  | GTAACATCCC | TTAAGGCATA | 2340 |
|    | TAAATTTTGT  | TCTGCAGGAG  | TTAATTTTCC | TGATTGTCTT  | ACAACTGCCA | CTTTATTTTC | 2400 |
|    | ATTAACCAAT  | TTACAAATG   | TTGCTTCATC | TATTTCAACA  | TGAAAACCAT | CAATTGCTTC | 2460 |
| 25 | TAATTTATCA  | ATCGTACCAC  | CTGTATGACC | TAATCCAAGC  | CCAATCATTT | TTGCAACAGG | 2520 |
|    | AACATCTACA  | GCTGCTACTA  | ATGGTGCTAA | AACCAATGTA  | GTTGTATCTC | CTACACCACC | 2580 |
| 30 | TGTTGAGTGC  | TTATCTACTT  | TGACACCTTT | AATATCACTC  | AAATCTATCA | TATCACCAGA | 2640 |
|    | ATTAACCATA  | GCCATCGTTA  | ATGCTGCACG | CTCATCATCA  | TTCATATCTT | GGAAATAAAT | 2700 |
|    | CGCCATTGCT  | AAACTTGATG  | CTTGGAATC  | AGGAATATCC  | CCTTTAACAT | AGCGGCCAAT | 2760 |
| 35 | AAAGAAATTA  | ATTTCITCCG  | TTGTTAGTGT | ATGACCGTCA  | CGCTTTTCTT | CAATAATGTC | 2820 |
|    | TATCATCTCT  | ATTTTATCA   | TCCTTTTCTT | AAAAAGCTTA  | GGACAAAGCA | TCTGCGCTTT | 2880 |
|    | CTCTAGTCCA  | TTTTTAAAG   | CACAAGCGAA | AATTATTATA  | GCAAGCTATC | GATTTCATTT | 2940 |
| 40 | TTAATATCAC  | AATTTCAGTG  | CGATGTTATT | ATTCTTAAAT  | AGATTGGTTA | TAACGTTAAA | 3000 |
|    | GTCCCTAATTA | AATTATCTTA  | GAATCATCAT | GGCATTTTATG | ATGCTTTAAA | GCTGATATCG | 3060 |
|    | ACATACTTAT  | ATATGGTTAC  | GATGTCCCAT | GCTTACATAT  | TTTTATAAAA | TTAGTAATCT | 3120 |
| 45 | GAATCTGCTT  | CTAAACCTTG  | CATAATTGGA | ACGCCTGCGC  | TGCACCAAT  | ACGTGTCGCA | 3180 |
|    | CTGCTTCAA   | CCATTTTATT  | GAAATCTTCT | AAATTACGTA  | CGCCACCTGA | TGCTTTTACT | 3240 |
| 50 | TCTATATCAG  | CACCTACTGT  | ATCTTTCATT | AATTTAACGT  | CTTCTGCAGT | CGCACCGCCA | 3300 |
|    | CTGCGAAAC   | CTGTTGAAGT  | TTTAACGAAG | TCCGCACCAG  | CGCTTTTGT  | TAATTCACTC | 3360 |
|    | GCTTTTACAA  | TTTCGTCAIG  | GTCCAACAAT | ACCGTCTCAA  | TAATCACTTT | TACTGTGTGA | 3420 |

55



EP 0 786 519 A2

|    |             |             |             |            |             |            |      |
|----|-------------|-------------|-------------|------------|-------------|------------|------|
|    | TTTAATGCGC  | CGATGTTGAT  | GACCATGTCA  | ATTTCATCTG | CACCAITTTG  | AATGCGATCT | 3540 |
|    | TCGTGTTCAA  | ATGCTTTTCG  | CGCAGTTGTC  | GATGCACCTA | ATGGGAATCC  | TATTACCGTA | 3600 |
| 5  | CAAACTAACA  | CCTCTGAATC  | AGCTAGTCGC  | TCTGCTGCAT | ATTTAACATG  | CGTTGGATTC | 3660 |
|    | ACACATAACG  | ATTTAAAAATG | GTATGCTTTT  | GCTTCATCGA | TGATTTGATC  | GATTTGCGTA | 3720 |
|    | CGTGTTGACT  | CAGGCTTCAA  | TAAAGTGTGA  | TCAATCAATT | TTGCACTATT  | CATTTTCTAT | 3780 |
| 10 | CTCCTCCTTT  | ATGGTTGATT  | ATAAAAAATAC | GGTTGTAAT  | TAGTTGATG   | AGCGTCAGGT | 3840 |
|    | TCATTTAAAT  | ATCAGGTTAG  | ATGTTTCGCTT | TTTATGTAAC | CGCATACATA  | TACTATTACA | 3900 |
|    | TTAATTCATT  | TCCCATAAAC  | AAACAATACA  | ATTGAACGTG | ATATCTTCAT  | TATGAACGAT | 3960 |
| 15 | GACTTGACAA  | CAAGCTAATC  | AGGATTATAT  | TTTTATAATT | CTTTAATTCT  | ATAGTACAAA | 4020 |
|    | AATTCGCAAA  | AAAGGGAAC   | AAATGTTATC  | TTAAAAATT  | TAATGAATAT  | TAAGGAGAAG | 4080 |
| 20 | ATAACAAATG  | ACAAAAGGTA  | CACCACATAT  | TCAACCAAT  | GGAGTAAAAA  | TGTGTAAC   | 4140 |
|    | AGTATTAAATG | CCTGCGATC   | CGCTACGTGC  | AAAATATATT | GCTGATAATT  | TTTTAGAAAA | 4200 |
|    | TGTTGAACAA  | TTTAACGATG  | TACGTAACAT  | GTTTGGTTAC | ACTGGTACAT  | ATAAAGGTAA | 4260 |
| 25 | AGAAGTTTCT  | GTAATGGGTT  | CTGGTATGGG  | TATTCCAAGT | ATTGGTATT   | ACTCATATGA | 4320 |
|    | GTTATACAAC  | TTCTTTGATG  | TAGATACAAT  | CATTGCTATC | GGTCTTTGTC  | GCGCATTACA | 4380 |
|    | AGAAAAATGTT | AACTTATACG  | ATGTTATTAT  | TGCACAAGCT | GCATCAACTA  | ATTCAAATTA | 4440 |
| 30 | TGTAGATCAA  | TACAATATTC  | CAGGTCATT   | CGCGCTATC  | GCTGACTTCG  | AGTTAGTAAC | 4500 |
|    | TAAAGCTAAA  | AATGTCGCTG  | ACCAAAATCG  | TGCTACTACA | CACGTAGGTA  | ACGTATTATC | 4560 |
|    | TTCTGATACA  | TTTTACAATG  | CCGATCCAAC  | ATTCAATGAT | GCTTGGAAAA  | AAATGGGTAT | 4620 |
| 35 | TTTAGGTATC  | GAAATGGAAT  | CAGCTGGTTT  | ATATTTAAAT | GCGATTCAATG | CTGTGAAAAA | 4680 |
|    | AGCACTTGGT  | ATTTTCACAG  | TAAGTGATCA  | TATTTTACGT | GACGAAGCTA  | CTACACCTGA | 4740 |
| 40 | AGAAGCTCAA  | AATTCATTTA  | CACAAATGAT  | GGAAATCGCT | TTAGAAATCG  | CAGAGTAACT | 4800 |
|    | TATTTAAATT  | GACTTTAATT  | GCTCTTTAAC  | AATGCGATTA | AACTCAAAAA  | GCCAACACAT | 4860 |
|    | TCGTGGCGTA  | TCCCATTTTA  | TGTGTTGGCT  | TTTATTTATA | TTATTACTTA  | TCTGTAGATT | 4920 |
| 45 | AGCTTAAGTA  | AGATTTAAAC  | ATCCAATTAT  | GTTTATCTAC | TGATGTTTGC  | ATACCTATAA | 4980 |
|    | ACATATCTTC  | TGATACATCA  | TGCGCAGCAT  | TACCAGCAAT | TTGATTGCG   | TTTCTAATT  | 5040 |
|    | GTTTGTAGAT  | ATTGTGGAAG  | TCTTGTGATA  | ATTCITCAAC | CATTGTTTCT  | GCAGAGTAAC | 5100 |
| 50 | CTTTCGCAGC  | TTCTTTAACA  | ATTGATTGCT  | CTAAGCATTC | AGTTAATGTA  | CCTACAGGGT | 5160 |
|    | TTCTCTCTAC  | CGCTAAAAAT  | CTTTCAGCTA  | ATTCGTCTAC | ATATTGGCTT  | GCTTCATTAT | 5220 |

55

AATTGTGTAG CTTTGTGTAA G

5301

(2) INFORMATION FOR SEQ ID NO: 444:

5

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1146 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 444:

|    |  |      |
|----|--|------|
| 15 | GAAGATGAAA GCTATATTGA AGATGATGTT ACTAAAAAGG AAGCTATTTT AAGTATGCAA  | 60   |
|    | ATTCTCTAAG GTTCTCTCA AAAATTAAAA GAGAACCGTT TAAAAGAAAC GATACAGTTA   | 120  |
|    | TATGGTAGAG ATGACTTTAT AGGTGGTATT GCTGTAGAAA TTGTTAGTAG TTCATTATAT  | 180  |
| 20 | GAGCAGCAAA TTCTTAACAT TATTTATGAA CACCTTGAGG ATATGAAACA GCATCAATCC  | 240  |
|    | ATCGATGCTA TCAACAAGTC CTATCATAAA CATACACCTG AATCTAAAT CAATTTGTG    | 300  |
|    | TCGCTTACTA AACAAGCACA ACACTCTATT TCAATTAGCT TAATCTTTGC GGTGATTTTA  | 360  |
| 25 | TTTGTTAGCG CTGTTCAAGT AGTACTTCAT TATCGTTTAA ACCAACAGC AGCATTGCAA   | 420  |
|    | CGATTATCAC AATATCATTT AAGCCGTTTC AAACATATATA GTACTTATGT AATGACACAT | 480  |
| 30 | ACGATTTTGT TATTGTTGGT ACTATTGGCA GTTAGTCTAT ATTTGTCTCA ACCACTCAGC  | 540  |
|    | TTAATATTTT ACTTAAATC ACTGTTACTT ATATTGATTT ATGAGATAGG TATCGTCTTT   | 600  |
|    | ATCTTATTC ATATTCAAAC AATAAGTCAT CGATTATTC TGACATTTAT ATATGCACCT    | 660  |
| 35 | GCTATGGSTA TCGTATACTT GATTATTTTC ATGTAAAGGA GCGTAACTGA TGATAGAAAT  | 720  |
|    | TAATAACCTT TCAAAGCGTT ACCGTAACAA ACAGATTTTC AATCATTTAA CTATGTCCTT  | 780  |
|    | TGATAGTAAT CGTTTAAACG TATTACTTGG TGATAATGGT GCTGGAAAT CAACATTACT   | 840  |
| 40 | TCGTATGATT GCTGGTATTG AAAAAGCTAA TGATGGAAC ATCAACTATT TCGGCGAAAA   | 900  |
|    | ATGGAATCAA AGACAATAC aAAATCAGT CGGTTATGTG CCACAAGACA TTGCGTTATT    | 960  |
|    | TGAACACATG ACAGTGGCTG AAAACATTAA ATTTTTTAAA TCACTTTGTA AAAATCCAAT  | 1020 |
| 45 | TAACGATACA ACTATCAACG AATATTTACA GCAATTAAAC TTTGATGATA CGTCTGCCAA  | 1080 |
|    | AGTATCTACA TTGTCCGCTG GGAATAAACG TAAAATTAAT ATATTAGTAG GTTTACTAGG  | 1140 |
|    | TCAACCTCGA ATTCTCATTT TAGATGAACC GACAGTTGGT ATTGATTTAA AATCTAGACA  | 1200 |
| 50 | TGACATCCAC CAACTACTTA ACATCATGAA ATCTAAATGT TTAATTATAT TAACTACCCA  | 1260 |
|    | TCATTAGAT GAAGTTGAAG CACTTGCAGA TGATATCAAG TTAATTGGCC AAGATCCCTT   | 1320 |

55

## EP 0 786 519 A2

|    |             |             |             |            |             |             |      |
|----|-------------|-------------|-------------|------------|-------------|-------------|------|
|    | CCCAAGCTGC  | GTATGATATC  | GCAACTTGGG  | ATTTTCTGTA | TTATCTACTT  | TGCAAGTATG  | 1440 |
|    | ACGTTGGGTC  | TACTGCATAT  | TGATTACCGA  | TGCCACCAGA | CATACGTTGG  | AAGTGATCGT  | 1500 |
| 5  | GAGGCGCTGT  | TGAATTACCC  | GTACTACCTG  | AATATGCAAT | TTGGTCACCA  | GCTTTGACTT  | 1560 |
|    | TATCACCAGC  | TGAAACAGTT  | AAACGATTAT  | TATGCATATA | CCATTGGTAG  | TTATTACTGT  | 1620 |
| 10 | TCGCTCTTTT  | AATCGTTACT  | TGATTGCCGC  | CACCATAGTT | ACTCCAACCT  | GCTTGTTACTA | 1680 |
|    | CTGTACCATC  | AGTTAATGAG  | TAACTGGTG   | AATTTTCAGG | CATTGCATAG  | TGCACACCGT  | 1740 |
|    | AATGCGCACC  | ACCACCGTGA  | TATTGTCCAT  | ATGTTGTAG  | TTGTTACGA   | CTTGTTAAAC  | 1800 |
| 15 | AGCTTGGCTC  | TTTCGCATGA  | CCACTAGcTG  | TCGCTTACT  | TGCTGATCCA  | CCATTTTGAT  | 1860 |
|    | TAGATGTGCC  | ATTAGGATAA  | TTGACCTTTC  | CATTACCATC | ATGGCTGTTA  | TACGCTTGGT  | 1920 |
|    | TGTTGTTACT  | ATGTGAATAA  | TAGCTCGCGT  | CTGGACCTAC | ATTGATTGA   | TAACCATATT  | 1980 |
| 20 | GATTAAATATG | CTGTGGCTT   | TGACTCGCTG  | TGTAGTCATT | GTTATCTCCT  | GCTGTGCTG   | 2040 |
|    | GATTACATA   | TGTTTGGCCG  | CTTCCATTGG  | CATTTGCATT | TTTTGGATAA  | CAGTTATAAA  | 2100 |
|    | AATAATGCGT  | ATGCTCTTGA  | GCATCTACGA  | ATGTATAGCT | ATAITCTTTA  | TTATCAAAAC  | 2160 |
| 25 | TTGCTTGATT  | CCAGTTACCA  | TCAGGTGTGT  | GATGATAATC | CCCATTAGAA  | TCAATTGTAT  | 2220 |
|    | AATAAGTACC  | ATAAGATACG  | TCTTGTGATT  | GTGTTGACAT | TTGTGTATGT  | GCTTGTGGG   | 2280 |
| 30 | TGTTTGTGCT  | TTCTGTGCA   | TCTGCTTGAT  | GCGCATTGT  | AAATGTAGCG  | AAcCCATCGT  | 2340 |
|    | TGCAATCGCT  | GCTGCTGTTA  | ATTTTTTCAT  | GTATAAAACA | TCCTCCATTA  | AAGTTAAAGT  | 2400 |
|    | TAGTTTTCAA  | TAAACTGTGA  | CTGCACATAC  | TAAAGAATT  | AGACAACCTGA | GTAAGGATT   | 2460 |
| 35 | TAATTTCTCAT | TTTCCAACCTA | TTTAATATTTC | CGAAATGTT  | TTACTAACT   | CATTACATTG  | 2520 |
|    | TCATTACAAA  | ATAGCCATAC  | ATTGATATTA  | AAATGACATC | TCTCACTGCA  | TTCGTTTAAC  | 2580 |
|    | CTTTTATAAA  | TTTTCAAAAT  | TAACAACTAA  | TCGTTCTGTC | ATGTTCTGCGA | TTCAACGCTA  | 2640 |
| 40 | ATGCATGATA  | GTAATCATCC  | ATCAAATCAT  | ATCAACCAAA | TTCCATTATC  | AATCGCTATT  | 2700 |
|    | GATTGTTCATT | CAACTTTCIA  | ATAGTGATAT  | GCTTCTCAGG | CTTAAAAATC  | GTCAATCTCT  | 2760 |
|    | TTCTATTAAT  | TAAATCATCT  | GTGAGCTTTA  | ATGCTACTAA | TTCAATTGCTG | CCATAATACT  | 2820 |
| 45 | TAATATATAA  | CGTCTTTGTA  | GTTAAATTTA  | TTACGGTCTG | ATACATCGTA  | TAGTGATTGT  | 2880 |
|    | CAGCATCATG  | CGGAGTACA   | ATTCCAATCG  | GTATATTTAC | CGCATCTAAT  | AAATAAAATG  | 2940 |
| 50 | CATTATTAA   | ATCCATTCT   | TTATCATTGT  | TTTGAGCAAT | GTTTGTCTTC  | ATAAATGCCA  | 3000 |
|    | TTCTCACAAA  | GCGCTCAGTT  | GAAGTAAATC  | CACCTGGCAA | TCCAATGTGA  | CCTGCTTCAT  | 3060 |
|    | TGCTTAAAGG  | TTCAATCGTT  | ACACCTTCCA  | ATAAATTTCG | TGTTGTCTGA  | TAAGGAGAAA  | 3120 |

EP 0 786 519 A2

|    |             |            |            |             |             |             |      |
|----|-------------|------------|------------|-------------|-------------|-------------|------|
|    | CACCAATAGG  | ATTATCTTTT | ATAACCACTT | CACCCCTCTT  | AAATGAAACT  | TCGACTGTAT  | 3240 |
|    | GTCCAGTTGC  | ATCGGAAACA | TGATAATGCA | ATGCGGGAAC  | TTACCCGATG  | TCATTTAAAT  | 3300 |
| 5  | ATACAGCTAC  | AACATGTATT | TGOGATGCTT | GTGTGTTTCT  | ATCTTCAATG  | CTTGTTGTAT  | 3360 |
|    | ATCCCAAAAT  | CCATGTCCAC | ATTTTCATTT | CGCTAATATT  | CATCGCGTCC  | GCTTTGTGTG  | 3420 |
|    | TTGATCCATA  | TGAATAATA  | CCTCGGAAAT | ATTGTGTTGA  | AATGGCAACG  | CcATGTTCAT  | 3480 |
| 10 | TAACACCATC  | ACCATAAATA | AAACCTTCCA | TATCTGTGTC  | TGTGCCAATA  | AAGCCATATT  | 3540 |
|    | GCGTTTGGCC  | TGTGTGCCCA | GTGCAAGATT | TCCAACGATA  | AITTTCTAGGC | GTCAGTCTGT  | 3600 |
| 15 | GCGAACCATC  | TAATGGATAA | TCATAATCCA | TCGTGCGTCC  | AAGAAGTACT  | TGATTATTTA  | 3660 |
|    | AAGTTTGTAT  | TGTGAATCCT | GTGCACATTG | TTCTCACTCC  | TCTGTACCTT  | CATTTACTTT  | 3720 |
|    | AATCACTTTC  | AAATAAGACT | GTTTCACTTA | AACATACTAT  | AAAAAATCAA  | TTATACAAGC  | 3780 |
| 20 | AATTAATGTA  | TATTCATTCT | CAATAACTGT | GGTATGATAT  | GTAAGGAAAT  | CATGACTTAT  | 3840 |
|    | GTGTGAGTGA  | ACGATCATCT | ATACATCCGT | TCACCTTCATC | TCATGACTTT  | CTATATTTAA  | 3900 |
|    | TTTTTACAAG  | GAGTGACATC | TGTGAATAAC | ACACAATCTT  | CACCACGCAG  | TAATATTATT  | 3960 |
| 25 | ATTGCGATTA  | TGTTGTCTGC | ATTAACATAT | TGGTTGTTTG  | CACAATCATT  | TATTATATATA | 4020 |
|    | GGACCTCTCG  | TTGGTCAAAC | ATATCAAACC | TCTCTGCCCG  | TGTTAAATTT  | ATCTATTAGT  | 4080 |
| 30 | TTAACTTCTC  | TCGCCACAGG | TATCTTCATG | TGCGTGCAG   | GTGATATTGC  | TGATAAAATA  | 4140 |
|    | GGACAACATGA | GAATGACATA | CATGGGTCTC | ATAATCAGTA  | TGTTTGCATC  | TCTTCTATTA  | 4200 |
|    | ATTATATCGG  | ACATCACTGC | ACTGCTCATC | ATCGGTAGAA  | TTTTACAAGG  | TCTATCAGCA  | 4260 |
| 35 | GCTATCTTGT  | TACCTTCAAC | AGTTGGCGTG | TTAAATAATC  | AATTTAAAGG  | AGAACATTTA  | 4320 |
|    | AGACGAGCGA  | TTAGTTATCT | AATGATTAGT | ACTGTTGGTG  | GCATCGGCTT  | AGCTGGTGTT  | 4380 |
|    | ATCGCGGGTT  | TAATGCGCTC | AAATTTCCGA | TGGCAAACGA  | ATTTTCATCAT | TAGTATAGTC  | 4440 |
| 40 | ATGCTTTTCA  | TTGCCATATT | GCTTCTAAAA | GGCACACCTG  | AAAAAGTAAG  | TCAACATAGC  | 4500 |
|    | CACCGTCATC  | CATTCGATTA | CAAAGGTATG | TCGATTTTCG  | CTGTTATGAT  | TGGTAGCTTT  | 4560 |
|    | ACATTATTGT  | TAACACAAGG | ATTCGAACAA | GGTTGGTTTA  | GTACATTTTC  | AATCATTTGT  | 4620 |
| 45 | CTGAGCATTT  | TTATCaTcAc | TACGTTGATA | TTCATCATCA  | TCGAACGTCG  | ACATGAAGTA  | 4680 |
|    | CCTTTTATTG  | ATTTCTCAGT | ATTACGCAAC | CGTCCGTTCA  | TTGGTGCATT  | TTTAAATAAC  | 4740 |
| 50 | TTGTGTTTAA  | ATAGCGGTCT | AGGCGTAACA | GTGGTCTTTT  | TCATATATGC  | TCAAACACAC  | 4800 |
|    | CTTGTTTATc  | CAGcTGCACA | ATCTGGACTT | GTTACATTGC  | CATATGCCAT  | TGTGCGAGTT  | 4860 |
|    | GCGATGATTC  | GTTTAGGTGA | AAAAGCAACA | TTACGTTTCG  | GTGGCAAATT  | GATGCTCATC  | 4920 |
| 55 |             |            |            |             |             |             |      |

## EP 0 786 519 A2

|    |   |      |
|----|---|------|
|    | TCACAATAIG TCATTGCAGT TATCATTGGT TTCGTCATAT GTGCGATAGG TAATGGTTTA | 5040 |
|    | GTGCGAACAC CTGGACTTAC GATTGCAATT TTCAGTATGC CTAATGAAAA AGTTGGTTTA | 5100 |
| 5  | GCTACAGGAT TATATAAAAT GAGTGGTACA TTAGGTGGCT CCTTTGGTAT AGCCTAAGT  | 5160 |
|    | ACTACAGTTT TCAGTATGTT ACAAATAAAC TATGCACCA GTGTAGCTGC AACCGTAACA  | 5220 |
|    | TTTATAGTCA GCATTGTATT GATGATCCTT GGCTCATGTT CTGCATACAT GATCATTCCA | 5280 |
| 10 | AAAACAGTTA AATCTTAAAT ATAATAGAAG AATTATGTTT CGAAATATCT TTATCACTTT | 5340 |
|    | AAAATGATAT ACAAGAAATC CAAGAAAAAT AAGCGAAGCT AATAAATAAA GATTCAATTA | 5400 |
|    | ACGCATCAGT ATTAGGATTC ACTCTAAAC GATTAATAGT TTTATAAGAA GGTGTTTGAT  | 5460 |
| 15 | CTTGAGCTAA CCACATCATT CGAATACTGT CATGAAGTAA TTTCTCTATT CTACGACCAG | 5520 |
|    | AAAATCAGA TTGAGTATAT GCATATAAGA TGAATTTTAA CATCATCTTT GGATGATAGG  | 5580 |
| 20 | ATGTTGCGCC ACGATGATGT CTGAATTCAT CGAATTTGCT ATCAGGTATC GTTTCACAA  | 5640 |
|    | TTTCATTAACT ATGTCGCGAA ATATCATTTT GAGGAATCTT AACAGAAGTT TTTATGGTA | 5700 |
|    | GTGTAAGTTG GGCAAGTGT CTTATTTTTT TAAAGTATTT CAAAGTAAAA TTACATGTTA  | 5760 |
| 25 | ATACGTAGTA TTAATGGCGA GACTCCTGAG GGAGCAGTGC CAGTCGAAGA CCGAGGCTGA | 5820 |
|    | GACGGCACCC TAGGAAAGCG AAGCCATTCA ATACGAAGTA TTGTATAAAT AGAGAACAGC | 5880 |
|    | AGTAAGATAT TTTCTAATTG AAAATTATCT TACTGCTGTT TTTTAGGGAT TTATGTCCCA | 5940 |
| 30 | GCCTCTTACT CTAATTATAT TCACTATCAA TTAGACAAAA TGGCCATTTT CAAATATCAC | 6000 |
|    | CGTTGTGTTT TGACCTTGAA TATATTTAAT ATAATCTCTT TTTTGAAAT CAGTTAACAT  | 6060 |
| 35 | TAATTTAGAT GTACCGTATT TTAACACTTT TTGCATTGTT TCTATTCTCA TTTTCTTAA  | 6120 |
|    | TAACCATCCA TCTTTTAACA CAATACGATT AACAGCATCA TATGATAATT CTACTGTTTC | 6180 |
|    | TTTAATTCA AATGTCCTGA ATGAAATAAT CGTGACATT AAAAACGTAT CACCAAGTA    | 6240 |
| 40 | ATAAACATCT AAATCATCAC GTTTATGTTG TCCAACTAAC AAACGACCAT ATTGGAATC  | 6300 |
|    | TTTTCTGGA TATTTCATTT CTAATAAACT AATAATCTCT TCTTCTTTA ATTGGAATTG   | 6360 |
|    | CATTTAAAAA CATCTCTCTT TAAGTTTTAA CAAGCCTTAA TTAATAATTT TTTCAATCAC | 6420 |
| 45 | ATAGTTCAAT ATACATCATT TCGTTATGTT TTTTAATACT TTGTTCAAAA ACAAATATTT | 6480 |
|    | TATCTTTTAA AATAAGACT TTTGTATTTT TAATCACAAT AACATTTTA AAATCTTGT    | 6540 |
| 50 | TATCATAATC ATTAAGAGGT ATTAACCTTA ATAATATATT CTCTCGTCTC AACCTTAATC | 6600 |
|    | GTATACCTCA GACGTCGTT TGTAGACAAT AAAAGTCATT CAGTCCTTCA TATGTCATCA  | 6660 |
|    | AATGTTTATC ATGATATGAT GAATATAATA ATCGGTATA TAACGTATG ATTAATTACA   | 6720 |

55

## EP 0 786 519 A2

|    |   |      |
|----|---|------|
|    | TGGTGTATCA GTTACAAC TGTCACATAT TTTAAATCAT AATGATAGTC GTTTTTCCGC   | 6840 |
|    | AACAACGATA AAAACGTAC ATGCTGTTTC AGAACGTTTA GGCTATGCCC CTAATAAACA  | 6900 |
| 5  | TGCAAAACAA TTGCGCGGCA GTAAAATCA AACTATTGGC GTCATTTTGC CTAGCTTAAC  | 6960 |
|    | AAATCCGTTT TTCTCAGCAC TGATGCAAG TATTCATGAC CATAAACCAT CTGATGTGA   | 7020 |
| 10 | TTTATGCTTT TTAACATCTA CAGCACTGA TTTGTATGAC AATATTAAAC ATTTAATTGA  | 7080 |
|    | TGGAGGTATT GACGGATTAA TTATCGCACA ATACATATCA TCCCCGGACG CCCTAAATAA | 7140 |
|    | CTATCTAAAG AAACATCATG TACCTTATGT CGTACTGGAT CAAATGACC ATCAAGGCTA  | 7200 |
| 15 | TACAGATTTT GTTCGGACAA ATGAATATCA AGGTGGACAA CTTCAGCAC AACATTAAAT  | 7260 |
|    | AGAAGCTCGT CACAACCATA TGATAATTGT TGCACCATAT GACATGATGG CGAATATGTC | 7320 |
|    | GACTCGTGTC GCTGGATTGG TCGATACTTT GCGCGGAAT CAATTGCCAG AACCACAAAT  | 7380 |
| 20 | CGTCCATACT GAATTATCTA AGCGCGGTGG GCTAACCATT GTTGATGACA TCATGGTTCA | 7440 |
|    | ATCTGCCACT GCAATCTTCG CTATTAAAGA TGAAGCTGCT ATTGGCATT TACGAGGACT  | 7500 |
|    | AATTGAACAT GGCATCAGTA TCCGAAAGA TATCTCATTA ATAGGTTATG ACGACATTGA  | 7560 |
| 25 | TTATGCAAGC TACGTCTCGC CACCTTTAAC TACTGTGGCA CAACCTATAA CTGATATTGG | 7620 |
|    | CAAAACATCT TTAACCTTAT TACTTCAACG ATTACAGCAC TTAGATAAAT CCATTGATAT | 7680 |
| 30 | GATTGAATTA CCAACGACTT TAAAAATTCG TGCAACAACG GGCTATCATC TTTCAAACTA | 7740 |
|    | ACTACGTATC TTCGAAATA TACTCATCAT TGTTAGGCCC TTAGCGTTGC TTTAATGCTG  | 7800 |
|    | AGGGTTTTTA ATCATAATTA TTTTACTAAG AAATTTAAAT AATAATGTAT GAATTTTTAA | 7860 |
| 35 | ATATGATTTA AACGTTTTCA GTTTTATGA AAACGCATGC ATTTTACAAA TAAAAATGAT  | 7920 |
|    | ACGATGGCAC TGGTAAACG TTTTACTAAA AACAAATCAT GAGGTGTATA ACATGAGCAT  | 7980 |
|    | TGTTGCATTA CTATCGGGT TAGGCCCTT AATTGGCTGG GGCTTCTTCC CAACAGTCGC   | 8040 |
| 40 | TTCAAAGTTT GGTGGTAAAC CTGTACATCA AATTATCGGT GCTACTGTAG GTACGTTAAT | 8100 |
|    | CTTCGTATT TTATTAGCGG TAGTCACATC AAGTGGCTTC CCTACTGGAA CCAATTGTCT  | 8160 |
|    | ATTCCGCTTA TTATCAGGTG CAGGATGGGG ATTCGGACAA ATCATTACAT TTAAGCGTT  | 8220 |
| 45 | CGAATTAGTC GGCTCATCTC GTGCCATGCC AGTCACAACA GCATTCCAAT TATTAGCGC  | 8280 |
|    | ATCTTTATGG GGTGCTTTG CATTAGGAAA TTGGCCAGGC ATTGGTCATA AAATCATTTG  | 8340 |
| 50 | ATTTACAGCT TTAGTCGTTA TTCTAATTGG AGCGCGTATG ACAGTTTGA GTGACGCAA   | 8400 |
|    | AGAAGCAAGT AACGCCAAAA ATTTACGTCG TGCAGTGGTA CTTCTGTAA TTGGTGAAT   | 8460 |
| 55 | TGGATACTGG TTATATTCAG CTGCACCGCA AGCAACTTCT ATTGATGGCC TAACTGCCTT | 8520 |

EP 0 786 519 A2

|    |             |            |             |             |             |             |       |
|----|-------------|------------|-------------|-------------|-------------|-------------|-------|
|    | AGCAGAGAAT  | CCATCCGTA  | ATAAAATTAC  | GTGGTTACAA  | ATTATTTTCAG | GTTTCTTCTT  | 8640  |
|    | TGCATTTTGGT | GCTTTAACAT | ATCTTATTTTC | AGCACAACTT  | AATATGAAATG | GTTTAGCAAC  | 8700  |
| 5  | TGGATTIAT   | CITTTCTCAA | CATCCGTTGT  | GCTTGCTACA  | TTAACTGGTA  | TTTATTTCTT  | 8760  |
|    | AAAACAACAT  | AAAACGTCAA | AAGAAATGGT  | TATTACAATC  | ATCGGCTTAG  | TACTCATTTT  | 8820  |
| 10 | AGTAGCCGCT  | TCTGTTACAG | TATTTATAAA  | ATAAGGAGTG  | TAGATGTCAT  | GA AAAATCA  | 8880  |
|    | GCTGTTTTAA  | ATGAACATAT | TTCAAAAGCA  | ATCGCGACAA  | TTGGTCATTT  | TGATTTATTA  | 8940  |
|    | ACGATTAAATG | ACGCTGGCAT | GCCAATTCCA  | AATGATCATC  | GTCGTATCGA  | CCTAGCTGTA  | 9000  |
| 15 | ACTAAAACAT  | TACCAAGCTT | TATTTGATGTC | TTAGCTACAG  | TGTTAGAAGA  | AATGGAAATC  | 9060  |
|    | CAAAAAATAT  | ACTTAGCAGA | AGAAATAAAA  | GAACATAACC  | CTACACAATT  | GCAACAAATT  | 9120  |
|    | AAACAATTGA  | TTTCATCGGA | AATCGAAATC  | ATTTTCATTC  | CTCACGAAGA  | AATGAAAGT   | 9180  |
| 20 | AACCTTAGCTC | ACCCATTAAA | TAAAGGTAAT  | ATTCGTACTG  | GTGAAaCAAC  | GCCCTACTCT  | 9240  |
|    | AATATTGCAT  | TAGAATCGAA | TGTTACTTTT  | TAAAAGTTAT  | AACCTGAAAG  | GAGCGTACAC  | 9300  |
|    | ATGACCAACA  | AAGTTGTTAT | TTTAGGTTCA  | ACGAATGTCG  | ATCAATTTTT  | AACAGTTGAA  | 9360  |
| 25 | AGATATGCAC  | AACCAAGCGA | AACATTACAT  | GTTGAAGAAG  | CACAAAAGC   | ATTCGGCGGA  | 9420  |
|    | GGTAAAGGTG  | CCAACAGGC  | TATTTGCCACT | GCACGCATGC  | AAGCAGACAC  | AACATTTATT  | 9480  |
| 30 | ACTAAAATTG  | GCATGATGG  | CGTTGCTGAT  | TTCATCTTAG  | AAGATTTTAA  | AGTAGCTCAT  | 9540  |
|    | ATTGATACAT  | CATATATTAT | CAAAACAGCT  | GAAGCAAAAA  | CGGGCCAAAG  | CITTTATCACT | 9600  |
|    | GTGAATGCAG  | AAGGACAAAA | CACCATCTAT  | GTTTATGGTG  | GTGCGAATAT  | GACGATGACA  | 9660  |
| 35 | CCTGAAGATG  | TTATTAACGC | AAAAGACGCT  | ATAATCAATG  | CAGACTTTGT  | CGTGCACAA   | 9720  |
|    | TTAGAAGTAC  | CCATCCCGGC | TATTATATCT  | GCAATTTGAAA | TTGCCAAGGC  | ACATGGTGTG  | 9780  |
|    | ACGACAGTAT  | TAAATCCTGC | ACCAGCGAAA  | GCATTACCTA  | ATGAATTAT   | ATCATTAAATC | 9840  |
| 40 | GATATTATTG  | TGCCAAACGA | AACAGAAGCC  | GAATTGTTAT  | CTGGGATTAA  | AGTAACTAAT  | 9900  |
|    | GAACAATCTA  | TGAAAGACAA | TGCCAATTAC  | TTTTTATCTA  | TAGGCATTAA  | GACTGTTTTG  | 9960  |
|    | ATTACGCTAG  | GTAAGCAAGG | TACATATTTT  | GCTACTAAAA  | ATCAAAAGCA  | ACACATCGAA  | 10020 |
| 45 | GCTTATAAAG  | TAAATGCGAT | TGATACAAC   | GCTGCAGGCG  | ACACATTTAT  | TGGTGCATTT  | 10080 |
|    | GTCAGTCGCT  | TAAACAAGTC | GCAAGATAAC  | TTAGCAGATG  | CTATTGATTT  | TGGTAATAAA  | 10140 |
| 50 | GCGAGCTCAC  | TCAGTGTACA | AAAACACGGC  | GCGCAAGCAT  | CTATTCTCTCT | ACTAGAAGAA  | 10200 |
|    | GTAATCAAG   | TTTAAATGAA | TCAAAACAGC  | CTATGATATG  | AAGGTTTAGC  | ATATAACATG  | 10260 |
|    | CAACATTCGT  | ATATCATGGC | TGTGCTTTTT  | TATCTTTATA  | AAACATCATC  | TATTAGAAT   | 10320 |

55

|            |            |            |            |            |            |     |
|------------|------------|------------|------------|------------|------------|-----|
| TTACATAGTT | AACACTAGTT | AATCTATTAG | TTAACATTAG | TTAATAATTA | GTTAATTTCC | 60  |
| ATTGTGATT  | TCATGTGATA | AATTTCTAAA | GCATACaATA | AATTTAATAT | GTAaaaAGAA | 120 |
| AGGGAATACA | CATGAAAAAT | AAATATATCT | CGAAGTTGCT | AGTTGGGGCA | GCAACAATTA | 180 |
| CGTTAGCTAC | AATGATTTCa | AATGGGGAAG | CAAAAGCGAG | TGAAAAACAG | CAACAAACTT | 240 |



EP 0 786 519 A2

|    |   |      |
|----|---|------|
|    | AAGTATTACA TCTAAAAGGT ATCAGAGAAG AACACGTAA CCAATACATC AAAACATTAC    | 360  |
|    | GCGAACACCC AGAACGTGCA CAAGAAGTAT TCTCTGAATC ACTTAAAGAC AGCAAGAACCC  | 420  |
| 5  | CAGACCGACG TGTGGCACAA CAAAACGCTT TTTCAATGT TCTTAAAAAT GATAACTTAA    | 480  |
|    | CTGAACAAGA AAAAAATAAT TACATTGCAC AAATTAAAGA AAACCCGTAT AGAGGCCAAC   | 540  |
|    | AAGTTTGGGT AGAATCAGTA CAATCTTCTA AAGCTAAAGA ACGTCAAAAT ATTGAAAAATG  | 600  |
| 10 | CGGATAAAGC AATTAAAGAT TTCCAAGATA ACAGAAGCACC ACAGATAAAA TCAGCAGCAT  | 660  |
|    | ATGAAGCTAA CTCAAAATTA CHTAAAGATT TACGTGATAA AAACAACCGC TTTGTAGAAA   | 720  |
| 15 | AAGTTTCAAT TGAAAAAGCA ATCGTTTCGT ATGATGAGCG TGTGAAATCA GCAAAATGATG  | 780  |
|    | CAATCTCAAA ATTAATAGAA AAAGATTCAA TTGAAAACAG ACGTTTAGCA CAACGTGAAG   | 840  |
|    | TTACAACAGC ACCTATGGAT GTAAAAGAGC ATTTACAGAA ACAATTAGAC GCATTAGTTG   | 900  |
| 20 | CTCAAAAGAA TGCTGAAAAG AAAGTGGCGC CAAAAGTTGA GGCTCCTCAA ATTCAATCAC   | 960  |
|    | CACAAATTGA AAAACCTAAA GTAGAATCAC CAAAAGTTGA AGTCCCTCAA ATTCAATCAC   | 1020 |
|    | CAAAAGTTGA GGTTCTCTCA TCTAAATTAT TAGGTTACTA CCAATCAITTA AAAGATTGAT  | 1080 |
| 25 | TTAACTATGG TTACAAGTAT TTAACAGATA CTTATAAAAG CTATAAGAA AATATGATA     | 1140 |
|    | CAGCAAAGTA CTACTATAAT ACGTACTATA AATACCAAGG TCGGATTGAT CAAACAGTAT   | 1200 |
| 30 | TAACAGTACT AGGTAGTGGT TCTAAATCTT ACATCCAACC ATTGAAAGTT GATGATAAAA   | 1260 |
|    | ACGGCTACTT AGCTAAATCA TATGCACAAG TAAGAAACTA TGTAACCTGAG TCAATCAATA  | 1320 |
|    | CTGGTAAAGT ATTATATACT TTCTACCAA ACCCAACATT AGTAAAAACA GCTATTAAAG    | 1380 |
| 35 | CTCAGAAAC TGCAATCATCA ATCAAAAATA CATTAAAGTAA TTTATTATCA TTCTGGAAT   | 1440 |
|    | AATCAATCAA AAATATCTTC TCTAGTTTTA CATCATTTTT TAAATAATTT TCGTAACAAA   | 1500 |
|    | CCGTGATTAA AAAGAACCCT TGATTCTCAA TCGAATCTAC GGTTCTTTTT TCATTTTCCA   | 1560 |
| 40 | TCAATTAAAT GCTTCTTCGC TATTGTGAG CCCACTTTTT TACCTGCAAC TTGTAAATA     | 1620 |
|    | ATCCTTACAT CGTTAACGAA TAGTTCATCA TTTAGTTGAA TCAGCTCAAC TTTATTAAT    | 1680 |
|    | TCATATTTTC ACAACTATT GCGCAATCCA TTCTTTTTC ACTACAAGCA CCATAATTAA     | 1740 |
| 45 | ACAACAATTC AATAAAATAA GACTTGCAAA GCATAGTTAT GTAGCTATAT AAACGCCCTGC  | 1800 |
|    | GACCAATAAA TCTTTTAAAC ATAACATAAT GCAAAAACAT CATTTAACAA TGCTAAAAAT   | 1860 |
|    | GTCTCTTCAA TACATGTTGA TAGTAATTAA CTTTAAACGA ACAGTTAATT CGAAAAACGCT  | 1920 |
| 50 | TACAAATGAA TTATTATATA TATGAACTTA AAATTAATAA GAAAGAAAGT GATTCTCATG   | 1980 |
|    | ATTAATAATAA AAATATTAAAC AGCAACTTTA GCAGTTGGTT TAATAGCCCC TTAGGCCAAT | 2040 |
| 55 |   |      |

EP 0 786 519 A2

CmTyCAAArG AcACaGACAT TACTAGCCAA CGATTTAGCT ATnACTCCAA ACCTTCCATT 2160  
GGATTTGGTA AAGGnT 2176

(2) INFORMATION FOR SEQ ID NO: 446:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1557 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 446:

|             |             |             |             |            |             |      |
|-------------|-------------|-------------|-------------|------------|-------------|------|
| AAAGCATGCG  | CTTAATATGAA | GTACGCTATA  | ACAAATTACA  | AGAACATGCT | ATTGTTATTGC | 60   |
| ATCCGGCACC  | TGTGAATAGA  | GGAGTAGAAA  | TACAAAGCGA  | TTTAGTAGAA | GCTTCAAAAT  | 120  |
| CAAGAAATTT  | TAAAGCAATG  | GAAATGTGCG  | TTTACTTAAG  | AATGGCAGTC | ATTGATGAAT  | 180  |
| TATTAATAAT  | GGTAAGGGGA  | CGAAAAATGAT | GAAATTAATT  | AAAAACGGTA | AAGTATTACA  | 240  |
| AAATGGCGAA  | TTACAACAAG  | CAGATATTTT  | AATTGATGGT  | AAGGTAATTA | AACAAATTGC  | 300  |
| ACCTGCAATT  | GAACCAAGCA  | ATGSGTGTGA  | CATCATAGAT  | GGGAAAGGTC | ACTTTGTGTC  | 360  |
| ACCTGGATT   | GTGCATGTTT  | ATGTTTCATT  | ACGTGAACCT  | GGTGTGGAAT | ATAAAGAGAC  | 420  |
| AATTGAAACT  | CGGTCTAAAG  | CTGCTCTGAT  | AGCGGGATT   | ACCACTGTAT | GTTCCAAATGC | 480  |
| TAAACACAGA  | CGGTACACAG  | ATGTCGTAGA  | ACATTTTGA   | GGTTTACAAA | AATTAATTCGA | 540  |
| TGACAAATGCT | CAAGTACGTG  | TATTACCTTA  | TGCTTCAATT  | ACAACACGTC | AATTAGGTAA  | 600  |
| AGAATTGGTT  | GATTTCCAG   | CACCTAGTAA  | AGAAGGTGCC  | TTTGGTTTA  | CAGATGACGG  | 660  |
| TGTAGGAGTA  | CAAACCTGCA  | GCATGATGTA  | TGAAGGCATG  | ATTGAAGCTG | CAAAAGTAAA  | 720  |
| CAAGCCATC   | GTAGCACACT  | GTGAAGATA   | TTCAATTAATC | TATGGTGGTG | CAATGCATGA  | 780  |
| AGGGAACGC   | AGTAAGAGAT  | TAGGTATACC  | AGGTATTCCA  | AACATTGTGT | AATCTGTGTC  | 840  |
| AATCGCAGA   | GATGTACTAT  | TAGCTGAAGC  | AGCAGGTGTG  | CATTATCATG | TATGTCATGT  | 900  |
| TTCTACTAAA  | GAAAGTGTTA  | GAGTCATTCG  | TGACGCTAAA  | CGCGCAGGCA | TTCAATGTTAC | 960  |
| AGCTGAAGTT  | ACACCAACACC | ATTTATTGTT  | AACAGAGAT   | GATATTCTCT | GTAATAATGC  | 1020 |
| CATTTATAAA  | ATGAATCCAC  | CATTGAGAAG  | TACTGAAGAT  | AGAGAGGGCT | TGTTAGAAGG  | 1080 |
| TTTACTAGAC  | GGATACAATT  | ACTGTATCGC  | AGACAGCAT   | GCACCACATG | CAGCGTATGA  | 1140 |
| AAAGGACACA  | CCGATGTGAA  | AAGCACCAT   | CGGAATGTT   | GGTAGTGAAA | CACGATCTCC  | 1200 |
| ATTATTATAT  | ACGCATTTTG  | TAAAAAATGG  | TGATTGGACA  | TTACAACAAT | TAGTAGATTA  | 1260 |



## EP 0 786 519 A2

|    |  |      |
|----|--|------|
|    | TAAATTTTAA TACTGCGGGG TGTCTTAAAA TGACATTTT AGTAACAGGG TTTGCGCCTT   | 1260 |
|    | TTGACAATCA AAATATCAAT CCCTCATGGG AAGCTGTGAC TCAACTAGAA GATATTATTG  | 1320 |
| 5  | GCACACATAC AATCGATAAA TTA AAACTAC CAACCTCTTT TAAGAAAGTA GATAATATTA | 1380 |
|    | TAAATAAAAC GTTGGCATCT AATCATTATG ATGTTGTACT AGCTATAGGA CAAGCTGGTG  | 1440 |
|    | GTAGAAATGC CATTACCCCA GAACGTGTG CCAITTAATAT TGATGATGCA CGTATTCCAG  | 1500 |
| 10 | ATAATGATGA TTTTCAACCT AITGATCAAG CCATTCACCT AGACGGTGGC CCAGCTTATT  | 1560 |
|    | TTTCAAATTT ACCaGTTAAA GCAATGACTC AAAGTATTAT TAATCAAGGA CTTCTGGAG   | 1620 |
|    | CACCTTCAAA TAGCGCAGGT ACATTGTGTT GTAATCACAC ACTTTATcAC TTAGGTTATT  | 1680 |
| 15 | TACAAGATAA GCATTACCTT CACCTACGAT TCGGATTTAT TCaTGTGCCA TACATACCAG  | 1740 |
|    | AGCAGGTcAT TGGTAAACCC GATACACCAT CTCATGhCCA TTGAGGAAAA GATnAGTTG   | 1799 |
| 20 | (2) INFORMATION FOR SEQ ID NO: 448:                                |      |
|    | (i) SEQUENCE CHARACTERISTICS:                                      |      |
|    | (A) LENGTH: 1341 base pairs  |      |
|    | (B) TYPE: nucleic acid   |      |
|    | (C) STRANDEDNESS: double   |      |
| 25 | (D) TOPOLOGY: linear   |      |
|    | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 448:                         |      |
| 30 | ACTTGGTTTT TTATTGTTTA TAAATAAAAC TCACTTAATA ATGTTTTTcAT AATCTTCTTC | 60   |
|    | GACTACTTAA TTCTTTAAGA TATTGTTGAA AAGAGACATT ACACTAGTTA ATTTTCAAAC  | 120  |
|    | AATACAAAAA GCGTCTACCT CCTACATATA ATTGTAGCGG AGATAGACGC TTAATATTTA  | 180  |
| 35 | TTTAAAAAAT ATTTTAAACC ACCGAATGTC ATAACATCAC GGGCAATCAT ACTTTCTTCA  | 240  |
|    | TCTGTTGGAA TAACGACAAC TTTAACTGGT GAATGAGGAT AGTTAATAAA TCCTTCTTTA  | 300  |
|    | CCAGTAGTA AGTTTTCATT TTTCTTAGGA TCCCAATAAA CACCATAAAA TTCTAAGCCT   | 360  |
| 40 | TCAAGAAGCT TCGCACGAAT TTCTACTGAG TTTTCCCGA TACCTGCTGT AAATACGATA   | 420  |
|    | ACATCAACAC CATGCATTCT CGCAGCATAT GATCCAATAT ATTTGTGAAT TTTAGAAGCA  | 480  |
| 45 | AATACATCTA AAGCCATTTG TGAACGTGCT TTACCTGATT CAGCTTCTTC TGATAAGTCA  | 540  |
|    | CGTAAATCAC TAGATGTACC TGATAATCCT AATAAACCTG ATTCTTTGTT TAAGATTTC   | 600  |
|    | AATACTTGTT CAGCAGTTTT ACCTGTTTTT TCCATAATAA ATGGAATTAA AGCAGGGTCA  | 660  |
| 50 | ATATTACCAG AACGAGTACC CATTGTTACA CCAGCAAGTG GTGTGAAGCC aTTGATGTAT  | 720  |
|    | CAATAGATTT ACCGCCATCG ATAGCTGCAA TTGATGCTCC ATTACCAATG TGACATGAAA  | 780  |
| 55 |  |      |

EP 0 786 519 A2

TATGGCTTGT ACCATGGAAA CCATACITAC GAATGCCATA ATCTTTATAA TAATGATATG 900  
 GCAAGCTATA TAGATATGCT TTTTCAGGCA TTGTTTGATG GAATGCTGTA TCAAAAAATTG 960  
 5 CCACATGAGG GATATTTGGT AATAATTAC GGAAGGCACG AATACCCATC AAGTTAGCTG 1020  
 GGTGTGGAAG CGGTGCTAAT TCGCTTAATT CTTCATTTTC CTTTCAACC TCATCAGTAA 1080  
 TAGCTACTGA TTCAGGGAAT TTTTACCAC CATGTACAAC ACGGTGACCT GTTCCATCGA 1140  
 10 TATCGTTAAT ATCATTAAATA ATAATGTGCG CTTTAAAAGC ATCCAACATG ATATCAACTG 1200  
 CCTCAACGTG ATCCTTGATA TCTGTACTG TTTTAACTTT TTCCCGGTG ACTTCAATTG 1260  
 TAAAAATTGA ATCCTTCAAT CCGATTCTTT CTACTAAACC TTTTGTACT AATTCCTCTT 1320  
 15 CAGGCATTCT AATAATTGA A 1341

(2) INFORMATION FOR SEQ ID NO: 449:

20 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1529 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 449:

TTTGAAGAT ACTACCGATG AAAATAGACA AAAGATTTTT CAATATTTAT CACCTGAAGA 60  
 30 AGTTGCAAAAT TTCTTTGATC AATTAGATAT TGATGACGAT GAATATGAGT TGCTATTTGA 120  
 TAAGATGAAT GCGACATACG CAAGTCACAT ATTAGAAGAA ATGTCATACG ACAATGCAGT 180  
 AGATATTTTA AATGAGTTGA CTAACCAAAA AGTTGCTAGT CTTTAAACAT TGATGAATAA 240  
 35 AGATGACGCG AATGAAATCA AAGCATTACT TCACTATGAT GAGGATACGG CCGGCGGTAT 300  
 TATGACGACG GAGTATTTAT CACTTAAAGC GCATACGCCCT GTTAAAGAAG CATTATATT 360  
 40 GGTCAAAGCG CAAGCACCAG ACGCAGAAAC AATATATGTT ATATTGTGCG TTGATGATGA 420  
 TGGTAAATTA GTAGGTGTTT TATCGCTAAG AGATTTAATT GTAGCTGAAA ATGATGCTTA 480  
 TATTGAAGAT ATTATGAATG AACGTGTCAT TAGTGTGAAT GTAGCAGACG ACCAAGAAGA 540  
 45 TGTTGCTCAA GTATGAGAG ACTATGATTT CATGGCTGTA CCTGTTATAG ATTACCAAGA 600  
 ACATTTGCTT GGTATCATCA CGATTGATGA TATTTTAGAC GTTATGGATG AAGAGGCTAG 660  
 TGAAGACTAC TTCTGTTTAA CCGGGGTATC AGATATCGAT TCGACTAATG ATTCAATCAT 720  
 50 TAAACAGACA TTAAACGTT TACCATGGTT GATTATTTTA ACATTTTATG GAATGATTAC 780  
 TGCACAATT TTAGGAGAT TCGAAAAAAC ATTAGAAAT GTAGCGCTAC TCGCAGCGTT 840

55

# EP 0 786 519 A2

|    |  |      |
|----|--|------|
|    | TCGTAACATT ACGACAGGGG AAATTAATGA GCAAAGTAAA TTTAGAATTG CATTAGAGA   | 960  |
|    | AGCAGGAAGT GGTGATTAT CCGGTGTTGT ATGTTCAACA ATATTATTTA CAATTATTGT   | 1020 |
| 5  | TGCAATATAT CATCAGCCAC TTTTAGCATT AATCGTTGCA GGAAGTTAA CTGTGCGAT    | 1080 |
|    | GACGTGGGG ACGTTTGTAG GTTCGATGAT TCCATTATTG ATGAATAAAT TAAATATCGA   | 1140 |
|    | TCCAGCAGTG GCTAGTGGAC CATTATTAC AACAAATTAAT GATATTATTA GTATGTTGAT  | 1200 |
| 10 | TTATTTTGGT TTAGCTACAT CATTATGCG GTTTTAACT CCAATAATTG TCAATCGATT    | 1260 |
|    | TTATCTTTA GTTATTGTG TTTTAGCAGC GTTTTAACT CCAATAATTG TCAATCGATT     | 1320 |
|    | AAATATTAAT TTCTTGCCAG TTGTTGTTGC AGAAATTTTG ATGGGGATTG TGATTGGAAA  | 1380 |
| 15 | TTCAATTTCTA AATATAGTAG AAAGGGATTC AATTCTAAAT ATTTTATCAA CGTTAGGCTT | 1440 |
|    | TATCTTTTAA ATGTTTAA GTGGTTTGA AATTGATTTT AAAGCTTTTA AAAAAGATAA     | 1500 |
| 20 | ACGGCAGCT CAAGACAAA ATGATGATG                                      | 1529 |

(2) INFORMATION FOR SEQ ID NO: 450:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1827 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 450:

|    |   |     |
|----|---|-----|
|    | TTCTGGAAC CAAAGTATTG TCATCTTCTA CTAGTAGTAT nGGCATCCAT ATCACCCTAAT | 60  |
|    | ATCATTAGT ATATTTCATA TTTTCTCTG ATTTTAATCG ACTTTGAAAA TCTTTAATCC   | 120 |
| 35 | GCGAGTCAAC TTCAAAGCCA TGAATCATCA TTTTTCGAT TGTGCTTGT ATAAAGTAAT   | 180 |
|    | AAATCGGCCA AAATAATCGA GGGATATAAT CGTATAGATG TATATAAAGC ACTGCGACT  | 240 |
| 40 | CTTTGATTAA TCTAAATCT AACTTCCCTT GATTAACTGT ATATTTTTC ACTAACTTC    | 300 |
|    | CACTCAATAA AATTAAAGTT ATTATTCCAT CAGCTGTTT TTCTATTTTA AATAGTCGA   | 360 |
|    | GCGGTGCGAC CTTATTCTTA ATATATATCT TAAATTGATC ATGTGATTTT TCTGTTTTC  | 420 |
| 45 | CAAAAGTTC TTTAGTGTAA CCCATCCATG CAATAAAATG GTTTACAACG TTCTTTAATG  | 480 |
|    | TCCATCCCTT TGGTAAACT ACCTTCATCG TTGATCTAAC ATCATCATAC TTGTAAACTT  | 540 |
|    | GTAATTCATC ATTAACATAA GAACGTTTAA AAACATAAAT TGTGTTTCT ACAGGTGTAC  | 600 |
| 50 | CATATGCACC TAGGCGTICA ATTGTTTCAT TATCATAACG ACTCCAGGT ATGTAGATTA  | 660 |
|    | CTTTTTTAC TTGATTGATA GCGCTGCTC GACCAAAAT ATCTGCTGCG ATTAATGTTA    | 720 |

(2) INFORMATION FOR SEQ ID NO: 451:

## 40

## 55





# EP 0 786 519 A2

(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 453:

|    |  |      |
|----|--|------|
|    | TCTCCAGTAG ACCTTGTGTA TGAACAGTTT CTTTCATATG AATGAACATC GTTTTAAAG   | 60   |
| 10 | TTTGTTCCAA CTCAGCCTTA TCAGGATAAT ATCTAGAGAC AGTCGTCTCT GGCATCTCCA  | 120  |
|    | TTGTATGATA TTAAACCTTA TGCAGCGACC CATGATTTTC GTTATAAACA AATGATGAT   | 180  |
|    | TCACTTCATC GAAATCATGA TCTTCTCTG CAATCCAAAA AACTGGTACT ACTTGTGTCT   | 240  |
| 15 | TATGTGTATC CGTTAATTC TTAGATAAAG TAATGATTGA AAATATTTTA TGGAAATGAT   | 300  |
|    | ACAATGGTCC CCCGAAAAGC CCGCTGTGTT GTCCACCAAT CACAACCTTT GAACCATTAG  | 360  |
|    | CTAAATGTTG TATGTTTAAT TCTTGTTCAC TTGAAAGCTT GAAAGCTTTA AATCACTCAT  | 420  |
| 20 | ATATTCACGA ATAACATCG CTAATGCCGC TTCTCTTCCA TTATTTTCTT TAGACATCCT   | 480  |
|    | TTTTTCAAAA CTAGTTTGTI GAGCTGCATC ATATTGAAAT AATCCTGTTA TTACAGGGTC  | 540  |
|    | ACTGCTCTTT ATTTTGGTA TAAACTGATC TTTTTCATTT AAACAACTA CTTTACAGTC    | 600  |
| 25 | CATGTTTTTT CTCCTTAAGT ACGCGATTAC AAACATAGT ATAAAGTCTA TACCGGTGAT   | 660  |
|    | TGCAATTTT ACGGCTTGAA AATCAATTTA ATCATGGAAA ATTTATAATA TTCATTGTTT   | 720  |
| 30 | TACATTTCGA AATCAATGAA AAACACAAGT GGTTTAATGT ATAATAATAG TAGTAAACAA  | 780  |
|    | ATAAGGGGTA GATAAATATG AGTGAAATCA AACGTCCTGA AATTAATTAC AAAACTGACG  | 840  |
|    | AATTATTGCA AAACCTTAGA GCGTTTGGCA ACAAGACTT ATACATGGTC AATGAGTTAA   | 900  |
| 35 | ACGGTCAAAT GATTGATGCA AGTTCAGATT CACCAITTTA TGGCATATTT GTCGaGATCA  | 960  |
|    | ATTAGGAGCT AGAATGGCAT TACTAAAAAA AGGTGATGTC GAAGAAATCT ACTTCCCAGA  | 1020 |
|    | TTTGAAGAT TATATATTAT TATGGAAGTT AGAAGTATTA CCAAAATATC AAAACAGAGG   | 1080 |
| 40 | GTACGCTTCA GAATTGATTG ATTTTGCAAA GAGTTTCAAT ATGCCAATTA AAGCCATTGG  | 1140 |
|    | CAGAAATGAT TCTAAGGATT TCTTTTTACA TCATGGATTT ACAGATGTGG AAGCTAAAAA  | 1200 |
|    | TATAGAGGGA CATGATGTCT TATTGTGGAA ACCATAAGAT AATAATATTC GACACTACGA  | 1260 |
| 45 | GCATGAAAAA GCATCTTTTC GTAGTGTCTT TTTTACAATT ACTTTCCTAA GCTAATATAA  | 1320 |
|    | GTAAATCATT TTCAATATTAT TTGTCTTAAC GTACAATATC ATTTAGTTGT TTCCATGTAT | 1380 |
|    | TAATTCATA ATCAGGTATA ATTCTGGAT TATGATCAAA TCCTCTAAAA TTAACACAGC    | 1440 |
| 50 | AAGTAGCTAT ACCCGCATTG ATTCCACCTA GAATGTCAGA TGTTAnAGAA TCTCCAACCTA | 1500 |
|    | TAATCGAGTG CTGCTTTTCA TCCTCACCAA TATCATTAAA AACATAATTA AAAAATTCGG  | 1560 |

55

## EP 0 786 519 A2

|    |            |            |             |             |            |             |      |
|----|------------|------------|-------------|-------------|------------|-------------|------|
|    | ACGGCGTCTG | ATTTAACCTT | CTCTTTTGGC  | TTTCGGTAC   | ACCATTAGTA | ACAATATATA  | 1680 |
|    | AATCATGTCT | TTTCGATAAT | TCGACAATTG  | TTTCTAATGT  | TTGATCAAAG | TATTTAACTT  | 1740 |
| 5  | TAGCTTCTGC | TAATCCATTT | CTAAATAACA  | CATCTGCACG  | ATGCCCATCA | ACTTCCATTT  | 1800 |
|    | GATGATGTTT | GAAGTAATTC | ACAAATCGTT  | CTGATAATAC  | TTGACACTTC | GKTAATTTAT  | 1860 |
|    | TTTGTGAA   | AGCTTCCCAA | TGTGGTGaT   | TGaTTTTTTT  | AAATGkTAAA | AAATCATCCY  | 1920 |
| 10 | TTGTTGCTTT | ATGATTAAAA | ACATTGCGCA  | TATAGTGAa   | CGCCCATTC  | TCTGcATCAT  | 1980 |
|    | AAAAATCAAC | AATTGTATCA | TCAAAGTCTA  | TCAAATATT   | TTTATATCCC | AATTTCCCCA  | 2040 |
|    | TCTCTATAT  | TGCTATGTA  | TCTAAATCTT  | AACAGAGGCT  | CAAATTTCTG | CAAAATAAAT  | 2100 |
| 15 | AAACTGAGTG | CATAACATTA | AAGTATGCTC  | ACCCAGTTTA  | TTTTAAAGAA | TATTAGTTAT  | 2160 |
|    | TATATTAGAA | TCCAAATAAT | TTACCTAGTA  | AACCCACACC  | GTTAGCAACG | ATGTCTACGA  | 2220 |
| 20 | TACTTGTGCC | TAATTTTACA | CTATCATGTT  | GTTGTGCAGC  | TTGCACAGTA | TTTGCAGATTG | 2280 |
|    | CTTCTGCTAG | TCCAGTCAIT | TAAATCTCTC  | CCTCACCTTT  | GAATAATAC  | TGATTACTTA  | 2340 |
|    | CATAACATAT | TGAAATTAGA | ATCCGAATAA  | TTTACCTAAT  | AAACCTACGC | CATTTTCAAC  | 2400 |
| 25 | GATGCTCACA | ATGCTTGTGC | CTAATTTTGC  | GCCATCAITTA | TTAATTGCTG | CAGTTACGGT  | 2460 |
|    | ATCTTTAAAT | GCSTTAAATA | AACCTTCCAT  | TGAAAACACT  | CCTTAAAAAT | TAAATTTGAA  | 2520 |
|    | GATAACAAAA | ACGTGCGTAg | YTTTTAAATC  | ACCGAAATGT  | TATTCGCTTA | ACGTTTGTGT  | 2580 |
| 30 | GTTGTTATTT | TAAAATAAAT | TTGATGCAAT  | TAGTTTGTGT  | ATCCGCACAA | CATCTTATAA  | 2640 |
|    | TGTACTTAAC | TGTATTTTAA | AGAGAAAAGA  | AATACAGTTA  | GGCATTCAAA | ACTGTATTTA  | 2700 |
|    | ACACAATTA  | GTTGCCTGAA | TTCTATTTTA  | AGCTTTATTG  | AACCTTTTTA | GATAAATAGC  | 2760 |
| 35 | TCTATAATAG | TGAAAAATAT | AAACATTTTT  | TATTTACAAG  | GTATTGCTAA | TTTAAGTTCA  | 2820 |
|    | TTTAGATATA | ATAATCTTGT | TGTTGTTAAA  | CGTGTCTG    | TAGCTCAGCT | GGATAGAGCA  | 2880 |
|    | ATGGCCTTCT | AAGCCATCGG | TCGGGGGTTT  | GAATCCCTCC  | CAGGACGTTT | ATAGGTATTT  | 2940 |
| 40 | TTATACGCAT | TACCAACAAA | AAGAGTTCCG  | TGATTACGGG  | GCTCTTTTTG | TTTTGAATTT  | 3000 |
|    | CAGTAATATA | GTATGATGCG | TCACCAAAAC  | GTCCCCCGCA  | TAGCCCCGCA | AAATACAGTA  | 3060 |
|    | ATTAAAACAA | GCATGCTTAT | TCGTTATAGA  | ATTTTTTGAC  | ACACAATTGA | CAGCGGTCTG  | 3120 |
| 45 | ACACTTGTTT | ATACATTTTT | AAATTAAGTAA | TTTGTGCTC   | AAATTTTATC | TATACTGCAC  | 3180 |
|    | CTGAACATCA | CCAACACTAC | ACCAAGATTT  | TTAACACTCA  | CCATTTGCAT | GCGTAGAGAT  | 3240 |
| 50 | TTTTATTATT | ATATTATTCC | TATAGATTTT  | GATACTATT   | AAAATTTTAT | GGACTTTTCA  | 3300 |
|    | GGGGCCCGAA | ATCCTATAAT | TATAATTATA  | TACATCTAAA  | AAAAATAACC | ACGTCATTCG  | 3360 |

55

# EP 0 786 519 A2

|    |  |      |
|----|--|------|
|    | ATACTATTGG CAAATTITATA AAGTAGTTCA GCGTTTTTCA ATGACATAIT GTCTAATGAT | 3480 |
|    | CTTTCATTIT TTCTCAITCT GTGTATTGTG CTTGTGGAA CTCCTGTTTG TTTCGATATA   | 3540 |
| 5  | TGTAAGCTGC TCAATCACT GTCTAATAGT TTTTGAATTT GATTTCATCAT TTTGTACACC  | 3600 |
|    | CCCTTGTAAG TCTTCAATCA TCATTTAAT TAAATAATTA CTTTCCACA TATTCCAAT     | 3660 |
|    | TTTAGGTTGC AAAGCATACC TCAATATCA TTAATTTTGA GATTTAAATG TCAATATATG   | 3720 |
| 10 | GTCTATCAAT CCAATATACA TACTCTAATA ACGTAATAGT ACACACTCTT CTTTATTAAT  | 3780 |
|    | GGCCATAGCT ATCATGATAT AATTAGTGAA GAAAATCACA TAAGAAAGGT TGTAAATCAT  | 3840 |
|    | GAGACTTCAA AAAGCACCTC TAGTAACGTC AGGACTAGTC TTAGGATTAT TAGGCCTGGG  | 3900 |
| 15 | TAATCTATTA AAAGACTTAT CTCTTACTTT AAACGCTGTT TGCGGAATCT TTGCTTTCTT  | 3960 |
|    | GATTGGATT CACCTTTTAT GTACTATGAT CAAATATTTT AATAATGTGA AAGAACAAT    | 4020 |
| 20 | AAACAGTCTC CTAGTTTCAT CAGTGTTCAC AACATTTTTC ATGTCTGGCT TTTTAGGTAC  | 4080 |
|    | TACTTATTTA AATACATTTT TTAGTAACAT AACTTTTATC AATAGCTTAA TAACGCCTAT  | 4140 |
|    | TTGGATTTTA TGCCCTGTGG GAATTATGAC GCATATGATT ATTTTTTCAA TAAATATTT   | 4200 |
| 25 | AAAAGATTTT TCACCTGAAA ATGTTTATCC TTCGTGAGCT GACTTTTTTA TTGGTATGn   | 4260 |
|    | TATCGCAGGA TTGACGGCAC CCGTTAGCGG ATATTTTTTC ATAGGTCAAT TAACAGTAAT  | 4320 |
|    | ATATGGCTTT GTAGCTACTT GTATTGTCTT ACCTATAGTT TTCAAGCGAT TAAAAGCATT  | 4380 |
| 30 | TCCATTGCAG ACGTCAATCA AACCGAACAC ATCGACAATT TGTGCACCAT TTTCTTTAGy  | 4440 |
|    | GGCTGCAGCA TATGTTATAG CTTTTCTTAA GCGCAATGCT TTTATCGTAA TTATATTTTT  | 4500 |
|    | ACTATTAGCT CAAATATTTT ATTTTATAT CATTATACAA TTGCCTAAAT TACTAAAAGA   | 4560 |
| 35 | ACCTTTTTCG CCGCTATTTT CAGCTTTCAC ATTCCCTTTA GTAATCTCAG CAACTGCTTT  | 4620 |
|    | AAAGAACAGT TTGCCTGTAC TTATGtKtCC AGACATTTGG AAAGGKCTTT TGTTTATCGA  | 4680 |
|    | AGTGTTATTA GCCACTGTAA TAGTACTTAG AGTCITTATA GGATATCTTC ACTTCTTTTT  | 4740 |
| 40 | AAAAAAGGAA AAACAAGATA AATTCTTnCG TAATGCGTCT CAGTAACACT ATTACCAAGA  | 4800 |
|    | ATTAACACGT ATATTTAATA  | 4820 |

(2) INFORMATION FOR SEQ ID NO: 454:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 4358 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

|    |             |            |             |             |            |             |      |
|----|-------------|------------|-------------|-------------|------------|-------------|------|
| 5  | ATTGGGATCT  | TTAATATCAC | CAATATTTTT  | AATATCTTCC  | GGATTCAATC | CATATACTTG  | 60   |
|    | TACTGTATCT  | GAGTATTTAA | TTGTGAAATA  | ATCACCTGAT  | TTAACTTTGT | CATCAACTGT  | 120  |
|    | AAATTGTGAT  | TTTAATGATA | AATAAATCTTG | GGCTGGTAGC  | ATTTTATTTG | TTTTTATCTGC | 180  |
|    | ATCAACGACA  | GTTAATGTTG | TATTTGATGT  | GATTAATACTA | TTAACATTTT | TAGCCTCTGT  | 240  |
| 10 | TGATGATGGC  | TGTACTGCTG | CTATACGCAT  | TCTTGTATTG  | AAACGTTTAG | GTGCTGTACT  | 300  |
|    | TTTTGGCAAA  | ATGATATCTG | CATTATTTTC  | ATTATTGTAA  | TTACTATTGT | TATCAACAAG  | 360  |
| 15 | AGTTTCATCA  | TTACTCTTGA | TAGCATACTG  | TTTGACATTT  | AATGTAGTTG | ATTCAGTTTT  | 420  |
|    | GGCATCTACC  | TTTTGGTTTT | CTCATTTAGT  | TTTGTAACA   | TTTACCAGCT | ATTATTTCTC  | 480  |
|    | TTGCAAACTA  | GTTTGTAAAG | CTTCTTGATT  | ACTTATAGTT  | TGTTTAGTGT | TTAAATCTTC  | 540  |
|    | ATTCGTAGAT  | TTTGGTGAAG | CTTGCTCATC  | TGATTTGGCA  | GTTGAAACTT | CAACTTTTAT  | 600  |
| 20 | TCCAGTGGTA  | GATTGTACAC | TTTCTTTTTC  | TATTAATTTA  | TTCCCAATTT | AAGTCGTTTC  | 660  |
|    | ATTACCTTGA  | GATGATACCA | TTTCTTTTGT  | ATTATCATTT  | TTAGTATTGT | CTTCTTGATT  | 720  |
| 25 | TAGTTGCTGC  | ATATCAACTT | TATCACTCGA  | TTGATTATCA  | CTTGCTGAAG | TTGTGCTGCTG | 780  |
|    | TTCAATTCTT  | TATTAGTACT | TTCTGCAGCC  | TTTGCTTCTT  | GGTCCCCAG  | ACCAAAAAAT  | 840  |
|    | AATGTTGTAC  | CTACTAAAAT | TGATGCTGTT  | CCCACGTGT   | ACTTTCTAAT | CGAAAAATTTA | 900  |
|    | TTTAACTGAT  | TGGATACCAT | GCCTTTCCCT  | GTTATTGCCG  | TTTTATTTTC | TCTGTTTAGC  | 960  |
| 30 | ATTAGATTAC  | TCCTAATTTA | TCAAATTTTT  | AAATAATACA  | ATTGTTTATA | ATACAAAAAT  | 1020 |
|    | GTATGATCAAT | ATAGTATTAT | ATTTTATAGT  | AAGSCACAAAT | ACTTTAATTA | TTTTTCTTTA  | 1080 |
| 35 | TCGTAAAAAG  | TATTTTAACA | TTTGTGTTTA  | AATAAAGGTT  | TTTATAGATT | TTGTAACTCT  | 1140 |
|    | TATTTAATCA  | TCATAAAAAA | TAGTATTATT  | TGCCCTTGAA  | ATTAATATCT | TAGCTTTTCT  | 1200 |
|    | AATTCATAGA  | CAATTAATT  | TCTGTAAACA  | ATTAAATTTG  | ATCTATTCTT | TAAAGATTTT  | 1260 |
|    | TTGTTTTATA  | TCTGGGAATT | TCTAAACAGA  | AAAAACCAGG  | CCACATGGAC | TCTGGTTAAGT | 1320 |
| 40 | TAATCATATT  | ATTTATTTTG | TTTTTTACGA  | CGACCGAATA  | ACAAATATGA | TCCTAATGCC  | 1380 |
|    | GCGAATAATC  | CACCGAATAA | TGTGCCAATTA | TTTGAATTAT  | TATTTTCACT | ACCTGTTTCT  | 1440 |
| 45 | GGTAATGCTT  | TAGCTGTTTT | ATGCTGATCT  | TTAACCGTAC  | TCATTGGTTT | AGCCCGGAGTA | 1500 |
|    | TGTTTACCTG  | CATCTGAATC | TGAATCGCTA  | TCTGAATCTG  | AGTCGTGTTG | TGAGTCGGAA  | 1560 |
|    | TCGCTATCTG  | AATCTGAGTC | GCTGTCTGAA  | TCTGAATCTG  | TATCCGAGTC | TGAGTCGCTA  | 1620 |
|    | TCTGAGTCTG  | AGTCGCTATC | TGAATCTGAA  | TCGCTGTCTG  | AGTCCTGATC | GCTATCTGAG  | 1680 |
| 50 | TCTGAATCGC  | TGTCCGAATC | TGCTGCGCTA  | TCTGAATCTG  | AATCGCTATC | TGAATCTGAG  | 1740 |

## EP 0 786 519 A2

|    |  |      |
|----|--|------|
|    | TCTGAATCTG AGTCGCTGTC TGAATCTGAA TCACGTCTGT AGTCTGAGTC GCTGTCTGAG  | 1860 |
|    | TCTGAATCGC TGTGAGAATC TGAGTCGCTA TCTGAGTCTG AATCTGAATC ACTGTCTGAG  | 1920 |
| 5  | TCCGAATCGC TATCTGAATC TGAATCGCTA TCTGAGTCTG AGTCGCTATC CGAATCTGAG  | 1980 |
|    | TCGCTATCTG AGTCTGAGTC GCTATCCGAG TCTGAATCGC TGTCTGAGTC TGAGTCGCTG  | 2040 |
| 10 | TCTGAATCTG AATCGCTATC TGAGTCTGAG TCGCTGTCTG AATCGCTGTC TGAATCTGAG  | 2100 |
|    | TCGCTATCTG AATCTGAGTC GCTATCTGAG TCTGAATCGC TGTGAGAATC TGAGTCGCTA  | 2160 |
|    | TCTGATGTTT CTCTTCGTA GTAGCCATTA TCAAGTGTGA AATCATCATG ATCCGTAATT   | 2220 |
| 15 | TTTACATCAA CTTCGCCACC ATCGGCATCT TTATCATCTT CAGTTGTATT TGTACCTGTT  | 2280 |
|    | TGAGTTAAGC CAGCAGGTTT TTCAAAGATA ACTTTGTATT TACCACTATC TAAATTTATCA | 2340 |
|    | AAGCGGTATT TACCATTTC ATCTGTYTCA GTTGTAACAA TTACTTCGCC TTTTTCGTTT   | 2400 |
| 20 | TGCAAGTAGA CTTTAACACC TTTAATTCCT TTTTCAGTCG AATCTGTGTT ACCATCTTTA  | 2460 |
|    | TTACTGTGCT ACCAAACATA ATCACCTAAA CTATATTTTG GTGTTTTGTA GAATCCACTA  | 2520 |
|    | TCTAATGTCA TGTGTGCAGC GTCTTTAATG ACACCTGTTG TAGTTAGTCC ATCAGAATCT  | 2580 |
| 25 | ACAGCATCAT CTGTACCTAC ATTGTGAGTT GTCGGTGTAT AACC GGCTGG TGTGAAAAC  | 2640 |
|    | TCTACACTAT AAGTTCATT GCTTAAACCA GTGAACGTAT ATTTACCATT TTCACTGTGT   | 2700 |
| 30 | GTGTCAGAT CTAATTCITT ACCGTTACTA TCTTTAAGAA TGACATAAAC ACCTTTAATC   | 2760 |
|    | CCTTTTTCAT TGGCATCTTG TTTACCATCT TTATTTGTAT CTTC CCATAC ATAGTCACCT | 2820 |
|    | AGATTATATT TCTTTTGCTC GCCATTAGCA GTTGATGAGC CATTACATT TGAATAACTA   | 2880 |
| 35 | TTTGACCAAC TATATTTAGT TTTGTCAGTG TCTAAAGTAT AATCAATTTT TCCATTATCT  | 2940 |
|    | GTTGAACTAT TATCTGGATA AGCAACTGTG TGAATGATGT ATTGTTTATT GCTGCTTGTT  | 3000 |
|    | TGGCCTTCA TTAAATCGAC TGTAGCTGTT TTATTATCAT TACTATAAAT AACATCGAAT   | 3060 |
| 40 | TGATCAGTAA CATCTTTAAG TTTTGAAGTA TCAGGGGTGA AACTATCCAC AAATGATT    | 3120 |
|    | TGATCTGTCA CTTCGTAAAT TTTGAAGTTT TTTGCATTTG GATTAAATTT ATATCCAGTT  | 3180 |
|    | AAATTAGTAA CAAAGCTTTG TTTAGTATAT GTATTTT TAG GTTGATTAC ATATGCAGTC  | 3240 |
| 45 | ATATTACGCG ATAAATCTTC ATTGTTAATA TAGTTTG TAC TTGAAATAAG CGGTTGTGCT | 3300 |
|    | TTTTTATTAC CATAATCGAC AATGATTCTT TCGCTATATG TATCATTACC TAAAGTTACT  | 3360 |
|    | TCCATTTTAT AAGCTGTTTT ATCAGTTGTT GCATTTT TAC GTTTCGCAAA TGCAACTTGT | 3420 |
| 50 | TCAAAGTAC CTCTAACATT TGTATATGTA TCTACATAGT TCGTAAAAGT ATATGTTGTT   | 3480 |
|    | GTGTTTGTG TACTATCATA AATACCTTTT GCAATAATAT TACCTTGGGC ATTATATAAA   | 3540 |
| 55 |  |      |

EP 0 786 519 A2

GTAAATGTAT CGCCCTCTTT AACAGAATCA TCGATTGTGT AATTGCTTT TAATTTTAAA 3660  
 ACATCACTTG AAGTGGCCA AAATTCAGTT TTACCAGTAG TCTGATTAAAC ATGTCCTTTA 3720  
 5 TCAATCGCAA TGTCAATATT TGA AAAAATGT ACTTTATCAT TAACATTGT TCCTTGTGT 3780  
 GGAGCTGCAA CAGTATTCAC TGCCATGCGA TTTAAAGTTC TTGGTTTAAT AGTCGTTGTT 3840  
 TTAGGTGTAG TTGAACATC TTTTGCTTGT GTTAAATTAC TTTTATCAGT TTCATTACTA 3900  
 10 TATGTAGTTG ATGATTATC ATTTGTTGTT ACATTGCTAG TTTTGTAGT AGATTGATTA 3960  
 GCTGTAGCGT TTTGTGGTGA TTGCATGTTA CTACTAGTTT CTTTAACTGT TGCACTATCA 4020  
 CTCATTGTCA CTTTAGGCTG ATCTGCAGTT GCAGTTTGGG TATTGTCTTT TAGTTGACGA 4080  
 15 CTATCAACTT TTTTAGTTGT TTTATCTICA CTGGGGCTG TCGTTTCATT TTTTGATTGA 4140  
 TTTAATCTC CATTCTATG TTCTGCCGCT TTAGCTTCAT GACCACTTAA CCAAAAAATC 4200  
 20 AATGTTGTCC CTACTAAAT TGAAGCAGTA CCTACAGAAT ACTTCTTAT CGAAAATTTG 4260  
 TTTAATCGAT TTGGTATCAT GCCTTTnCTA TTTGThGCTG TCTTTTATA ATTCATTTAA 4320  
 TAATACTCCT TTAAATATC AAAATTTGAT AAATATAA 4358  
 25 (2) INFORMATION FOR SEQ ID NO: 455:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1060 base pairs  
 (B) TYPE: nucleic acid  
 30 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 455:  
 35 TTGACTTCTT AATTGAGCAT TTTCTGCACT TAATGCTTTG TTTCTTTTAA TAAGTTGCTT 60  
 TCTTGATAAA ACTTCGGTAT CTATTTTACT ATTACTATAC CTTTGATTTA AAACATAATAT 120  
 40 ACCAATTAAAT GCTACAATGA TAATGATAAG TACAACATAA AAAGACATTT TTTCAACAAAT 180  
 CCTTTTGTAC TTCTTTAACT TTGTATACAA TAATAATTAA TAAAGATTAA TTGTATTTCA 240  
 ATTTCCACAA TTTTATTAG TTGATTTTAG TTCATCATTG TTATAATCAA ATTATAAACT 300  
 45 GACAGATATT GATGTTCAAT GAATATGACG TGAAGATTTC GTGAATTCAA GTTTATGTCG 360  
 AATTATGTTT ATAACGGTCA TTTAAATGAC AGAATTAGTG CACTCATAGT ATTTTGAAGA 420  
 TTGAATTCAT TAATTTTAAA ATGTATAATG ATATTGTGA AAGCGCTTGC TTAGGAGGTG 480  
 50 TATTTGAGAG TGAATGAAAT GAATGCTAAA GAACAATTAG TGGACAATTT AATGAAAAA 540  
 TCATCGCAAT TATTTAAATT TCACGGTGAA GTTGCCATGC AGCTTTTCTT AAATGATGAA 600



TATAACGAAA AAGGTTGGTT AAATGAATCA GGTGCCAAAT ACCCTGTTGA AGTTGCCATT 1080  
 TTAAGAGATA ATGTATTATT GACTATCGAC ACATCAGGTT CTGGTTTGAA CAGACGTGGT 1140  
 5 CATTGATTAG CACAAGGTGA AGCACCAATT AAAGAAACGT TGGCAGCAAG TTTAATCCGT 1200  
 TTGTGCCAAG TGGGAAAGGT GATTACACCT TThAATTGGT CCCATTTCGG GTTCnGGTTA 1260  
 CA 1262

10 (2) INFORMATION FOR SEQ ID NO: 457:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1142 base pairs  
 (B) TYPE: nucleic acid  
 15 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 457:

20 CCTGGCTGCT TTATCAGCAT CTACTACTTT AAAACCGAAT ACGGATAAGA GTTCGATAC 60  
 TGTGTATTTT CCTGAGGCGA TTCCACCTGT TAGACCAATA ACTTTCGGCA TAAITTCAC 120  
 25 CTTTCTTTAT TTTTGACATA CTGACAATA ATGACTATTT CTGTGCGGCA TGATTTTGT 180  
 TTCAAITTTGA CTTCCACACA CTTTGACATC CGGCTGCTTA TATACATTAA GATGCAATTG 240  
 CATCTCACCA GTTTTTCCAT CAGCATGACG ATAATCTGAA ATACTGTGAC CGCCATATTT 300  
 30 AATACCTTCT TCTAGTACTT CTCTAACATA ATAAAAAACC ATTTCTTGT GTTGGTGTGT 360  
 TAAGTCTTTT ACTTTTTTAT CTGTAAAAAC ACCTGCACGA AACAACGCTT CACATGCGTA 420  
 AATATTTCCA CAACCTGCGA TTACTTTATG ATCCAAAATC ACTTGTITGA TTGGTTTATT 480  
 35 CTTATTAGAC TGTGTAGTAA TTCGATTTAA ATAATACGTC AATGCTTCAT TTGAAAAAGG 540  
 TTCAGGCGCT ATTTCTAAAA ATGAAGGATA AGATGCTACA GACGCAACAT TTCTAATTTT 600  
 TCCAAAACGA CGTATATCTG AATAAATTAA CTTTTGTICA TTGACAACAT CAAAAATAAC 660  
 40 ATGCCAATGC TTACGATAAT TAGGTATCAT AATATCTTCA AGTTCATCTA CAATGAAAAA 720  
 ACCGCCGCCC ATACCTAAAT GACTAATTAA TGTACGTTGT TCTCGTTTAT TATCTAGCTG 780  
 45 AAAAAACGATA TATTACTTCT TTCGTTCTAC ATTTGTAATG GTATAGCCTT CCGATAAAGT 840  
 TTTAAAAGTA TCTAATTCAA TTCCTTTTAT AATTGTTTCC TTGCTTTGAG CTTTACCTTC 900  
 GATTACTTTA TCCGAAAAATA TAACGTGTTC AATTTTTTGA TTTATAACGT AGGGTTCAAT 960  
 50 TCTCTTTTTT ACATGTTCTA CTTCGTGTAA TTCGGGATA CCATTAACCT CACTTTATTT 1020  
 TGCATCATAC CAGGTGCAC CATAACTTGA GTCTACTTTT AATGGAACAT CTAATTGCAA 1080

55



TT

1142

(2) INFORMATION FOR SEQ ID NO: 458:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1814 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO: 458:

|    |  |      |
|----|--|------|
| 15 | CCTTTAGTAA ACAATCCTTC TTTAGTTTAA GTACGTTGTT CCAATCCAAA TAATTGTAT   | 60   |
|    | TTCAATGCCT CGCCGATTG AGTCCGCTA AAGTTATCAT CTTTCATGTT AGGCGTGTG     | 120  |
|    | GTAAACATGT GTATATCACT GTTTAAACGG TCTTTTAAAG CTTGCGTACC TTGTACATCG  | 180  |
| 20 | TATTGCTTAT AAATATAACC ACCATCAACA GAGCCTTCTG TTTCTCTACC TTCGCTATCA  | 240  |
|    | GCATAAACAG TCGGTTCTAA AAACAACACG TTAGCTTCTC TTTGTTTCTC AACTTCTACA  | 300  |
|    | GGATCTAAAT TTAATATACC TTTAATAAGT AACATAGCGT CATTTAAATC ACTCATATAG  | 360  |
| 25 | TTAGCAsymy CTGATTCAGC ATTATCATAC AAATCAATTA AAGTGATTAC TTTCTCATAA  | 420  |
|    | TCCCCTTTTC TTTCTTCGTT GTTGCTAAAT TCTGTAATAG GCATACGTTT GAAAGAGTGT  | 480  |
|    | GATTCAAAAC CGTTTTTCAG TGGTGTGAGC TTCAATCCAT TTGTTCTACT GGTAAAGATAT | 540  |
| 30 | CTATAAACAC CGTGTGAAGT GAATAAATCA ACTGTAAACA CTTCACTCTC GTCAGTCTTG  | 600  |
|    | TCTATTGTTT TAGTTCCTAA ATATCTAACG CCTGCGATAC TATTACGTTT AATTGTATTG  | 660  |
|    | TCGTATATGA CAAAAGTACT CATTGCATCA CTCCTGTATA AACGCGTTTC ATCATCTTGG  | 720  |
| 35 | TTTCTAATCA TTAATCATA AGCTTTGCCA TAAATTGACA AATCTAATCC TAAAGACTTA   | 780  |
|    | TTGTGTGACT CAACATCATT TAAATCATTG AACGCTTCAA TAACTTCTAA TACATCTTTG  | 840  |
|    | TCATCATCTT GATATTGAAT TGGATTACCC AAGAAATAGC CGTTGATAAA ATCGCTAATA  | 900  |
| 40 | TAAGATGCGT AATCATGCGC TACACGGTTA TCTGCCATGT ACTCTTCTTT GCGTCGTGTT  | 960  |
|    | AACTCAACTA AGTTCTTAGT TTTACCTTCG TAATAATCAC TTAACACTTT CAATCTAGGT  | 1020 |
| 45 | CGTTGGTAAT CCAATGTGAT TTCAATGTAT TTAATTACTT CATTAACGTT TTGTAATAAA  | 1080 |
|    | TCCGATTCCG TCCCGTCATA TGTGTAAACA ACATTGGCTT CATCAITAAA TAAGTAATTT  | 1140 |
|    | ATGTTTCCCC GTAGATCTGT ATCTGTTTCA AATTCGTTTA CTTTTAACAT TTGTTCCCTC  | 1200 |
| 50 | CTATAATCCT AGAGATTTTA TTGTGTCAAC TTTCGAAGTG AGATTGTGTC GTTTTCTAAC  | 1260 |
|    | CGGTCTGTAG AATCGTTCCA CTGAATAACG CAACGAATCG ATACAATGAT TGTATGTATC  | 1320 |

CTCTTCAATA GTCTTGAAAC AACGTTTCATC AACAATGATT TCAAAATGCA TTAAGAATTG 1440  
 TAACCCCTTGT ACAACCGAGC CCTTCCCTTT TTTGGTTGGT AAAATCCTTT TAAGCCCTAG 1500  
 5 ATTCCCTTAAT TCAGCTATAC TTTTITGTTT TGCATATCT GCTGTAAATT CTTCITTTAGC 1560  
 ATAACCAAGT TGCTTTATGA CATTAGCTAT TTCATCATTC AGCATACCTT GTTTAAACATA 1620  
 CTCTTCAAGT ATGTATAACT TCITTTTCTT TACATCTATT TTAGAATGTA TAAAAGCACT 1680  
 10 AGGATCATT ACGTAGCCAA AGTCCAATCC AAAATAAGAA GGTAAATGTC TTAACATC 1740  
 TTTATTATT AAACGTTTTT CATACTTAGG GAAACCAAT TTGTCTAGTG TAGCAAAATC 1800  
 ACCTAACGCA TAAA 1814

## (2) INFORMATION FOR SEQ ID NO: 459:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 686 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 459:

AATTAGATT ATTACCCCTCC TTAAAAATA TTGTTTCAC AATTITTTAT TACCTATTTA 60  
 CTGGTTTCAT GTCTTATGGG GCATTTTACT TGTGGTATT TGAAAATGTG CAAAATTTAA 120  
 30 TCTTATATGT TTCTTGGCTT TTCATGACTA TGCTATTTAT GTTTATGAAT ATGCATTCAA 180  
 TTATAGATAA AAAAGTACAT ATATTCTTAA AGTCTAATAA ATAGTTACAA ATTTAGTTAG 240  
 TTTTCAATTG TTAATTAGGG GTGGTAAACA GTGCTTTGTG AATCTAGACA AATTTATAAA 300  
 35 AATCCTAAAT ATCGAGTTAT TAGATATAAT AATGAATATT TCATGGTCGA TTTAGTAAGT 360  
 ACTTGGATTA CTTATTTTIT CCCTATGATT AATTGGTTTT TGCCCCAAAA ATACGCAAAA 420  
 ATTAGCGAAA ATGAATTGA AAGGTTAAAT ATAGTCGAGC CTGTTAAAAA TAATGTTTTT 480  
 40 TGGCCGGTGT CAGGAAGTTC AGTTCTATTT GGAATTATAT TGAGAAAGTA CGGTAACTTC 540  
 TTTAATGTTT AGTTTGAAAA ACAACTAGCA ATCACTGTAT TTTTATCAT GTTAATAGGG 600  
 45 ATGTTAATTT TTTATTTTAA TCTAAATAAA AAATTAACAT TAAAAATTTT TAATACCAAC 660  
 GTGGGTAATA AGAATAGGAG TTGTAT 686

## (2) INFORMATION FOR SEQ ID NO: 460:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1300 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double

EP 0 786 519 A2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 460:

|    |  |      |
|----|--|------|
| 5  | ATCTGCAATT ATGGGCACAC CCAAGCTTAT GCAAGTAACT AAAGGAGAAG TACTTTTAGA  | 60   |
|    | CGGTGTAAT ATTTAGAAAT TAGAAGTTGA TGAAAGASCA AAAGCAGGAT TATTCTTGGC   | 120  |
|    | AATGCAATAT CCATCAGAAA TTACAGGTGT TACAATGCT GATTTCATGC GTTCAGCAAT   | 180  |
| 10 | CAATGCGAAA CGTGAGAAG GACAAGAAAT CAACCTAATG CAATTTATTA AGAAATTAGA   | 240  |
|    | TAAAAACATG GATTTTCTAG ACATAGATAA AGACATGGCA CAACGTTATT TAAATGAAG   | 300  |
|    | TTTCTCAGGT GGAGAGAAGA AACGTAACGA AATCTTACAA TTAATGATGT TAGAACCTAA  | 360  |
| 15 | GTTTGCAATC TTAGATGAAA TCGATTGAGG GTTAGACATC GATGCATTAA AAGTTGTATC  | 420  |
|    | TAAAGGTATT AACCAATGC GTGGGAAAA CTTTGGTSCA TTAATGATTA CACACTATCA    | 480  |
|    | ACGATTATTA AATTACATTA CTCCTGATAA AGTACATGTA ATGTATGCTG GTAAAGTCGT  | 540  |
| 20 | TAAATCTGGT GGTCCAGAAAT TAGCAAAAAC TCTTGAAGAA GAAGGATATG AATGGGTTAA | 600  |
|    | AGAAGAGTTC GGTTCAGCTG AATAATCTTA TTAATACAGT ATCCATGAGA TGTTTCATCTA | 660  |
| 25 | TATATGATGA AAATGAACAT TTATACGAAA TAGTAAATTT CATCAAGTAG GAGGAAAAAG  | 720  |
|    | TTATGACAAAC TGATATTTTG TACATTCTG AAGAACCACT TGTGTGTTAT TCTAAAGCCC  | 780  |
|    | ACAATGAACC TTCTTGGATG ACAGAAATTAC GTAAAAAAGC TTTGAAATTA ACAGAAACTT | 840  |
| 30 | TAGAAATGCC AAAACCTGAT AAAACAAAAT TAAGAAAATG GGATTTTGAT TCTTTTAAAC  | 900  |
|    | AACACGATGT AAAAGGTGAT GTTTATCAAT CTTTATCACA ATTACCTGAG TCAGTAAGAG  | 960  |
|    | AAATTATTGA CGTAGATCAT TCTAAAAACT TAGTAATTC ACATAATAAT ACGATTGCGT   | 1020 |
| 35 | ACACACAAGT TGATGATAAT GCATCGAAAG ATGGCGTTAT CGTTGAAGGT TTAGCAGACG  | 1080 |
|    | CTCTTAGTAA CCATAGTGAT TTAGTACAAA AGTACTTTAT GAAAGATGCA GTAACAGTAG  | 1140 |
|    | ATGAACATCG TATCACACG CTACACACGG CATTAGTTAA TGGTGGCGTA TTTGTTTATG   | 1200 |
| 40 | TTCTTAAAAA TGTAGTTGTA GAACATCCAG TACAATACGT TGTGTTGCAC GACGACGAAA  | 1260 |
|    | ATGCAAGCTT TTATAACCAT GTTATCATCG TTACTGAAGA                        | 1300 |

45 (2) INFORMATION FOR SEQ ID NO: 461:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1115 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

55

EP 0 786 519 A2

|    |             |             |            |            |             |            |      |
|----|-------------|-------------|------------|------------|-------------|------------|------|
|    | GACAGCAAA   | TCAAGATACA  | TAAATAAAG  | ATATTGTCAT | AAAGCATATT  | CTGTTCGAG  | 60   |
|    | ATAAATCTAA  | AAAGAGACCG  | CTAAATATTA | GCCAATATCA | TTTAACTGAA  | GATGTTAATG | 120  |
| 5  | AAATTTTAA   | TGATGATTCA  | TTAGATATTA | TCGTGAAGT  | CATGGGAGGA  | ATTGAACCAA | 180  |
|    | CTGTGATTG   | GTTAAGAAACA | GCACTTAAAA | ATAAAAAACA | TGTTATTACC  | GCAAAATAAG | 240  |
|    | ATTIATATAGC | AGTACATCTT  | AAACTTTTAG | AAGATTTAGC | AGAAGAAAA   | GGTGTAGCTT | 300  |
| 10 | TAAAGTTTGA  | AGCGAGTGTA  | GCAGGTGGTA | TTCCGATCGT | AAATGCCATA  | AATAATGGTT | 360  |
|    | TGAATGCGAA  | TAATATTTC   | AAATTTATGG | GAATTTTAAA | TGGTACCTCT  | AATTTTATTT | 420  |
|    | TATCTAAAA   | GACTAAAGAG  | CAACGACAT  | TTGAGGAAGC | ACTTGATGAA  | GCGAAAGAC  | 480  |
| 15 | TTGGTTTTGC  | TGAACCGGAT  | CCAAGTATG  | ATGTAGAAGG | GGTAGATGCA  | GCGCGTAAG  | 540  |
|    | TTGTCAATTAC | ATCATATTTA  | TCAATTAACC | AAGTCATTAA | ATTAAACGAC  | GTTAAACGAA | 600  |
| 20 | GAGGAATTAG  | TGGCTAACT   | TTAACTGATA | TTAATGTAGC | CGATCAACTG  | GCGTATAAAA | 660  |
|    | TTAAATTGAT  | TGTAAGGGA   | ATATATGAAA | ATGGCAAAGT | TAATGTCATG  | GTAGAACCAA | 720  |
|    | CGTTAAATTGA | TAAAAGCAT   | CAATTAGCAG | CTGTAGAGGA | TGAATATAAC  | GCGATTTATG | 780  |
| 25 | TTATTGGTGA  | TGCCGTTGGT  | GACACGATGT | TTTATGGAAA | AGGAGCAGGC  | AGTTTAGCAA | 840  |
|    | CAGGTAGTGC  | CGTTGTGAGT  | GATTTATTGA | ATGTAGCATT | ATTCCTTGAA  | TCAGATTTAC | 900  |
|    | ACACATTGCC  | ACCACATTTT  | GAATTAAGA  | CAGATAAAAC | ACGGGAAATG  | ATGGATTCAG | 960  |
| 30 | ATGCAGAAAT  | TAATATTAAA  | GAAAAATCCA | ATTTCTTTGT | AGTAGTGAAT  | CATGTCAAAG | 1020 |
|    | GTTCAATTGA  | AAATTTTGAA  | AATGAGTTAA | AGGCAATATT | ACCATTTCAC  | CGATCATTA  | 1080 |
|    | GAGTTGCAAA  | TTACGATAAT  | CAATCATATG | CCGCTGTTAT | AGTTGSAATTG | GAATCATCAC | 1140 |
| 35 | CGGAAGAATT  | AATCACTAAG  | CATGGATACG | AATTGACAAA | GTATACCCAG  | TAGAAGGAGT | 1200 |
|    | TTAATTATAA  | TGAGAAGATG  | GCAAGGATTA | GTAGAAGAGT | TTAAAGCACA  | TTTACCAGTA | 1260 |
| 40 | AATGAAAATA  | CACCAAAATT  | AACATTGAAC | GAGGGAAATA | CACCACTCAT  | TCATTGTGAA | 1320 |
|    | AATATGTCTA  | AAATACTAGG  | CATAGATTTA | TATGTGAAGT | ATGAAGGTGC  | CAATCCGACA | 1380 |
|    | GTTCAATTTAA | AGATCGCGGT  | ATGGTAATGG | CTGTGACAAA | AGCAAAAGAG  | CAAGGTAAAG | 1440 |
| 45 | AAATGTGAAT  | ATGCGCTTCG  | ACTGGAATA  | CATCAGCGTC | TGCAGCAGCA  | TATGCAGCGA | 1500 |
|    | GAGCAGGTTT  | AAAAGCTATC  | GTCGTAATAC | CAGAAAGTAA | AATTGCATTA  | GTTAAATTGT | 1560 |
|    | CGCAAGCAGT  | AATGTATGGT  | GCAGAAATCG | TTTCTATTGA | AGGAAACTTT  | GATGAAGCTT | 1620 |
| 50 | TAGAAATTGT  | AAAAGAAATT  | GCAAAAAGTG | GCGAAATCGA | GCTTGTAAAC  | TCTGTCAATC | 1680 |
|    | CATTAGAAT   | CGAAGGACAA  | AAGACAGGCT | CATTTGAAAT | TGTACAACAA  | TTAGACGGTG | 1740 |

55

EP 0 786 519 A2

|    |  |      |
|----|--|------|
|    | AAGGCTTTAA AGAATATCAT GAAGCTAAAG GATCACAATT GCCGAAAATG TTTGGCTTCC  | 1860 |
|    | AAGCTGAAGG CGCATCACCA ATTGTTCAAA ATAAAGTCAT TAAAAATCCT GAAACGATTG  | 1920 |
| 5  | CAACTGCTAT TCGAATTGGT AATCCTGCTA GTTGGGATAA GGCGACTAAT GCTCTTAAAG  | 1980 |
|    | AATCAAATGG ATTAATAGAT AGTGTTACTG ATGATGAAAT TCTAGAAGCA TATCAGTTAA  | 2040 |
|    | TGACAACTAA AGAAGGTGTC TTAGTGAAC CAGCGAGTAA TGCTTCTATT GCAGGTTTAA   | 2100 |
| 10 | TTAAATGCA TAGACAAGGT AATTACCTC AAGGTAAAAA AGTAGTTGCT ATTTTAACTG    | 2160 |
|    | GTAATGGATT AAAAGATCCT GATACTGCTA TTCACTACT AGATAATCCG ATAAAGCCAT   | 2220 |
| 15 | TGCCAAATGA TAAAGATAGC ATTATCGATT ATATTAAAGG AGCTTTATAA CATGTCGAAT  | 2280 |
|    | GTTTTGGAGT TAACAATTCC TGCATCAACA GCCAACCTTG GAGTTGGCTT TGATTCTATA  | 2340 |
|    | GGTATGGCTT TAGATAAAIT TTTCATCTG TCTGTAAAGG AAACATCAGG GACAAAATGG   | 2400 |
| 20 | GAATATATTT TCCATGATGA TGCATCTAAG CAATTGCCTA CTGACGAAAC AAACCTTTATT | 2460 |
|    | TATCATGTAG CACAACAAGT TGCTTCTAAA TATAGTGTG ACTTGCTTAA TTTATGTATC   | 2520 |
|    | GAAATGAGAA GTGATATTCC ATGGCAAGA GGGTTAGGTT CGTCAGCTTC TGCTTTAGTA   | 2580 |
| 25 | GGAGCTATAT ATATCGCAA TTTATTTGGT GATATCCAAC TGTCTAAACA TGAGGTATTA   | 2640 |
|    | CAATTAGCGA CTGAAATCGA AGGACATCCT GATAATGTTG CGCGACCAT TTATGGTGGT   | 2700 |
|    | TTAATCGCTG GATATTATAA TGATGTCTCG AAGAAACGT CAGTtGCACA TATCGACATA   | 2760 |
| 30 | CCAGACGCTG ATGTGATTGT AACGATACCA ACTTATGAAC TAAAAACAGA AGCATCAAGA  | 2820 |
|    | CGTGCTTTAC CACAAAAAAT AACACATAGT GAAGCGGTTA AAAGTAGTGC AATTAGTAAT  | 2880 |
|    | ACAATGATTI tGgCATTAGC ACAGCACAAT TATGAATTAG CAGGTAAACT CATGCAACAA  | 2940 |
| 35 | GATGCGCTTC ATGAACCGTA TCGTCAGCAT TTAATIGCTG AATTTGATGA AGTGAACAA   | 3000 |
|    | ATTGCTAGTC AACATAATGC CTATGCAACT GTAATTAGTG GTGCTGGACC AACTATTTTA  | 3060 |
| 40 | ATATTTAGTC GTAAAGAAAA TAGTGGGGAA TTGGTTCGCT CTTTAAATAG TCAGGTAGTA  | 3120 |
|    | TCATGCCAAT CTGAA   | 3135 |

(2) INFORMATION FOR SEQ ID NO: 462:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1209 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 462:

# EP 0 786 519 A2

AGhTCAATAT TTAGATCAAG CCGTTTTAAG TAATTACGAA CAAGTTTATA TCATTTCATGG 120  
TAAAGGTACA GGTGCACCTC AAAAAGGTGT ACAACAACAT TTGAAAAAGC ATAAAAAGTGT 180  
5 TAGTGACTTT AGAGGTGGTA TGCCAAGCGA AGGTGGATTT GCGGTTACCG TTGCAACACT 240  
AAAATAAATT ATAATTGAT AAATTAATA GCTGCAGTTA AAATAATGTA AAGCAACAAG 300  
AATACATTTC AAACATGTTA TTGAAATAA GCATAAAAAA TGAGCAAAAT GAAATACATG 360  
10 AAGCATGITA TCTGATATAA TTGAACATC ATAATAATAA TTAAGGAGGA TTGGCATTTA 420  
TGGCAATCGT AAAAGTAACA GATGCAGATT TTGATTCAAA AGTAGAATCT GGTGTACAAC 480  
TAGTAGATTT TTGGGCAACA TGGTGTTGGT CATGTAAAAA GATCGCTCCG GTATTAGAAG 540  
15 AATTAGCAGC TGACTATGAA GGTAAGCTG ACATTTTAA AATTAGATGT GATGAAAATC 600  
CATCAATGC AGCTAAATAT GAAGTGATGA GTATTCCAAC ATTAATCGTC TTTAAAGACG 660  
GTCAACCAAT TGATAAAGT GTGTGTTTCC AACCAAAAGA AAAGTTAGCT GAAGTTTATG 720  
20 ATAAACATTI ATAAGTTACA ACCAATGACG ACTGGGGCAT TTCTTTAATG AATTGCTCCA 780  
GTTTTTGTTI GTGTTTTTAA TATAAAAAGT TGAATGATAA GTCATCATAT TGTTTACGAC 840  
25 TTGAGAATGG TGGGATTAAT AAATCTATGA ACGTTAAATG ATAATCTAGC ATGCTGATAG 900  
ATTTGTAGCA GTTGGTTTGA TAAACCATG TTCAATATTA CATGATGTGC ATGAAAAGTC 960  
ATACTCGAAG ATGTTGATTA TTAAGTAGAA TTAGTGGTGA TAAATTGAA GCACCTTTTGT 1020  
30 AGCATCATTI ATTTTAAAT TAGAAGGGGG GATATTTTGG GAAGACTATA AGCAACGAAT 1080  
TAAAAATAAA TTAAATGTCG TACCTATGGA ACCAGGATGC TATTTAATGA AAGATCGTAA 1140  
TGATCAAGTG ATATATGTTG GCAAAGCTAA AAAGCTAAGA AATCGATTGC GATCATATTT 1200  
35 TCACGGGTG 1209

## (2) INFORMATION FOR SEQ ID NO: 463:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 2410 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 463:

AGTTCACAC GACGAATTTT ACCTGAGTTT GTTTTTGGTA AGTCGTCAAC GAATTCATTC 60  
50 TCTCTGGAT ATTATATGG TGCAACTTCA TTTTAAACAA ATTGTTGTAG TTCTTTAACT 120  
AACGTATCAT CACCCGCAGT ATGGTCCTGT AAAATAACGA ATGCTTTAAC AATATTTCCT 180

55

EP 0 786 519 A2

|    |             |            |            |            |             |             |      |
|----|-------------|------------|------------|------------|-------------|-------------|------|
|    | GCATCTTCAA  | CTTCAAAAGG | CCCAATCGTA | TAGCCTGAAC | TAATAATAAT  | GTCATCTCGA  | 300  |
|    | CGTCCTTCAA  | ACCAGAAATA | ACCATCATCA | TCTACATGAG | CTAAGTCACC  | AGTAGTGATG  | 360  |
| 5  | TATTCACCTG  | TTTGCGCTTT | CGCCGTACGT | CTGCGCTCTT | LATAATACCC  | TTTGAAAAGT  | 420  |
|    | GCTGGCAAAAT | CAAGTGGTAC | TGCAATATTC | CCTTTCGTAT | TAGCAGGTAC  | GCTATTCCCC  | 480  |
| 10 | TCATCATCTA  | CTACAGTGAC | CGAACTACCC | GGAATGCTT  | TACCCATTGA  | TCCAATCCTC  | 540  |
|    | TGTGGTGAT   | CTTTTAAAAA | GCCTATAAGC | AAGGTACTTT | CAGTCTGGCC  | ATATCCATCT  | 600  |
|    | CTTACAGTTA  | AATTAAGTA  | TTTCTTGAAT | TGTTCAACTA | CTTCTCGATT  | TAGTGGCTCA  | 660  |
| 15 | CCTGCAGAAA  | CGGCACTATG | TAAATGCGTT | AAGTCATAAT | CATTTAAGTT  | CTGTAAATTA  | 720  |
|    | GCCATCATAC  | GATATTCTGT | CGGTGTACAA | CATAAAACAT | TAATTTGATA  | TTTTTGGAAGC | 780  |
|    | AAITCTAAGT  | ATGTTTCAGG | ACTGAACCTT | CCAATAAATA | CAAAAGCAGT  | TGCACCTGAA  | 840  |
| 20 | CCTAATACAG  | ATAAGAAAGG | ACTCCATACC | CATTTTGGCC | AACCTGGTGc  | TGCTGTGTGCC | 900  |
|    | CAAACTAAGT  | CATCTTCATT | aATACATAaC | CAATGTTTGT | GTGCCATTGT  | TaAATGTGcA  | 960  |
|    | AATCCCCaTc  | CATGACAATG | TGTAACGGCT | TTAGGATTGC | CAGTTGTACC  | AGATGTATAT  | 1020 |
| 25 | GACAGAAATC  | CCATATCATC | ACGCGTCGTA | TCTGCCATTT | CTAGTTTGT   | ACTTGCCTTT  | 1080 |
|    | TCTTTTTTCAG | CTTCAAGTGA | AATCCATCCA | TCTTTTTGAC | CGGCAATAAC  | AAATTTAGTT  | 1140 |
|    | AACCATCAT   | ATTCTTTAAT | TTTTTCAAAT | TCAACTGTGA | ATGGCTCTAG  | TGCAATAAAT  | 1200 |
| 30 | GCATTAATTT  | CACCATGTGT | GATACGGTAT | TGTAAATCTT | TAGTTCTTAG  | CATTTCAGAA  | 1260 |
|    | CATGGAATGA  | TTGCAACACC | TAATTTTAAA | GCAGCAATAT | ATAATTGATA  | CGTCGCAATA  | 1320 |
|    | GATCGTGCCA  | TCATAATGAG | TACTTTATCG | CCTTTAGATA | AACCGTGCGA  | TGCTAAAACA  | 1380 |
| 35 | TTACCTACTT  | TATTAGACTG | TTCAATGAGC | TGTTGGTAAG | TGACTGATAT  | ATCTTCGCCT  | 1440 |
|    | TCAGTATTAT  | GATATAAAAT | TGCCCTTTTA | TCTGGTATGT | GGCTATATTT  | TTCGATTTC   | 1500 |
| 40 | GAAATAATGT  | TATATTTTTC | AGGCGCGAAT | AGAGCTGACT | TTTGCAATAAC | TAACCTCCTT  | 1560 |
|    | TCATACATCC  | ACTTTTCCTG | TGATGAACAT | TGTAATTTTA | TAAATGAATT  | ATATACATCA  | 1620 |
|    | TACGCTATC   | TTTACAGAAT | TTTCAATTAA | ATAGGGTTAA | ATACCAAAGT  | CCTGCACACT  | 1680 |
| 45 | ACACTTTGAC  | ATGACGTAGC | ATTCAAGGAC | TTTCAAATGA | TTGAGGGTTG  | ATATCTCGGG  | 1740 |
|    | CTAGACCATA  | TCAGCTAATT | CAATACGAAT | ATTGTATGAT | AATTCACGAT  | TAATTTATTTT | 1800 |
|    | TACATCTGCA  | CCTTTCGAAG | TGCCACGATG | CTTGTGTGTA | TGCTGTGACT  | CAGCTGAATT  | 1860 |
| 50 | TGCGCAATGA  | TAAATGTCTT | GCCTATTTTC | CCACAGCGTA | ATAATGATAT  | AGTGTCTTACC | 1920 |
|    | AGCTGTCTTA  | GGCTTAAAAA | ACCTTAATGC | TTTAAATCCA | TCAACGTTT   | TTAAATGCTT  | 1980 |

EP 0 786 519 A2

ATTTAACACA CATAATGAAT CATTTGATAA ATCATTATC GCTTCTAGCA CATCGTAATA 2100  
TGCAGTGICA TTATTTTTT GTATTGTGAG ACAATCATCC AGTTCTTCTA TTACATAACT 2160  
5 TCTATATTCA TCATAAATTT TCATAATAAA TGCCTTCATT TCATTTATAT TTTTGGTCAT 2220  
ATTACTKTAT ATCTATTACT AATKCAATCC CGTATTATAT AATTACAATC ATAGTTTGGC 2280  
TyCTTTTTAA AAGATAAGAC TTTGTAAAAA GTATTAAATAT TTCATGCAAA TGGGGGACAG 2340  
10 GAGTCGCCCA CTATTTTGT GTCTCAATT TCATGATCAT TAITTAACAT TAGTCATGAA 2400  
AATAGCGGAC 2410

15 (2) INFORMATION FOR SEQ ID NO: 464:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 590 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
20 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 464:

25 TTTATTAATT GTAAAAAATT GAGTAAATTA TCTTTACATT CTAATTAGT CTTAACTACA 60  
ACTAGTATCC TAATAATTAT AGGAGCTATT ACATTCTTTT TATTAGAACA GTTTAATACT 120  
ATGCAACATA TGGGACTAGT TGAAAAAATC GGAAATTCCT TTTTCCAATC AGTAACAACA 180  
30 CGAACAGCGG GTTTTAACAG TATAGATATA GCAAGCATT rCAAAATCTAC CGCATLaaTG 240  
TTAATGCTAC TTATGTTTAT TGGTGGTGCC CCTCTCAGTG CAGCTGGAGG AATTAAAAAT 300  
ACTACTTTTG CAGTTCGTT TATTTTGTGA CTAAATTATA KACGTAAAGA AAATAATGTT 360  
35 TCAGTATTCA ATAAAGAAAT ATCTGACAAA CATATAAAAC TATCTATTGT TACCATTAAAT 420  
ATCTCATTC TATTATCAG CATCATTACT TTTATATTAT CGATAATTAA TCCGAACATA 480  
TCATTAATCA AGTATTATT CGAAGTGGTT TCTGCATTCG GAACAGTAGG GTTAAGTAGT 540  
40 AACCTTACCA CAGAATATCA TGGTATTACT AAAATAATTA TTATATTCGT 590

(2) INFORMATION FOR SEQ ID NO: 465:

- 45 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 905 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
50 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 465:



# EP 0 786 519 A2

TATTCGATT GATCTTTTA AAATTATGTT TAAAAACATG TCTAATGATT CCGCATCATT 120  
 TTTTGACCA TCAATAAGCG TTTCAGCAA CCCCTTAATT GAAGTAATAG GTGTTTTTAA 180  
 5 TTTATGTGAA ACATTGCTA CAAATTCAGT TCTTAGATT TCAAGTGTG TCAGATTTGT 240  
 TATATCATGC ATCACAATA AAATCCCTTG CAAACTTTTT TGAGACCTAG TTAATAATCG 300  
 AAGCATGAA ATATCAAGT ACTTGCGATG GACTTGGTTT ATTGCAACTT CCAATTGTTC 360  
 10 ATAAATAGGT TTTTCAACTT TAAACCTTC TAAATTAAT TGCTCAATT CAGTATTAAC 420  
 ATAGCCGTGA TAGCTACTT GTTCAATATT ATGCGAGATG TTGAACGTG CATAATACG 480  
 TTTATTTGCA ACAAGATT TTTCAATTG ATCTATCATT AAAATAGCAC TTGAATATT 540  
 15 TTCAATCGTT GTTTTTAAAC GGTGGGATTG AATTTTTTGC TCATTATTA GCTTTTGAAG 600  
 GCGTCTGCT AAATCATTTG TAGACACAAA AAGCGCTTTA GTTCTACAA CATTACTTTC 660  
 20 AGGTACAGT ATGTGATAAT AACCATTTC CAACAATTG GTTGCAATAG TAATCTCTTG 720  
 AATGGGACGG ATTAATGTAC GCTTAAACT ACGGCTTGA AAATACAGAC AAATGAGTAC 780  
 AACTAAACAT GTCAAAATAA GATATTTC CAACGTCAA TGCATTCTG TAATATCGTT 840  
 25 ATTGTAACCT TTAATCCATA CATGATAACC GTTAACCTTC TTATAAAAA TAAAAACGTC 900  
 CCTTT 905

## (2) INFORMATION FOR SEQ ID NO: 466:

30 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1016 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

35

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 466:

TTTGGTTAGC CCATAAAGA AAAAAACAAG TAGTCATTTT TAAACAACAT ATCAAGTCTA 60  
 40 CCCAAGAAAT ACGTTTGTAC AAAGCGAAAG TGCTTGAACA CAAAGATGAA ATAGCAAATT 120  
 TTATTTCTTT CGAACCACAA AGTTTGAAT TTTATTATTT TACAGAAATG GAATTTTCAG 180  
 45 AAGAACAATT AAATGAAGTT TCGCCAATTA GAATTAATTT CAATGTATA AGACACACA 240  
 AAGATTTGAT AAAGCATATG CGAATATAT TTTTGGCTAG ACTTATTTCA GAAGATAATG 300  
 ATAAAAAGAC ATATATGTTT TATAAACGCA AAGTATTAA CGATAACTTT TTAGATAAAT 360  
 50 ATATGCAGAA ATTTTCACCG GCAACATACA CAATAATATT TGTAATGTC TTAATATGGT 420  
 TATGTATGAT TTTATATTTA AATAATTTTT CGGATGTAAA ATTATTAGAT GTTGGCGGGT 480

55

## EP 0 786 519 A2

ATTTTAGTTT TGAACATATA CTTATGAATA TGCTTTCATT ATTTATTTTT GGTAAAAATG 600  
 TCGAAGCAAT TATTTGTTCA TGGCGGATGT TAACGTATA CTTTATTGCA GGGTTGTTG 660  
 5 GAAACTTTGT ATCACTATCA TTTAATACGA CTACAATTC AGTTGGGGCT AGTGGTGCTA 720  
 TATTGGTCT GATTGGATCA ATTTTTCGA TGATGTATGT TTCAAAAACA TTTAACAAAA 780  
 AAATGTTAGG ACAGTTAITA ATTGCATTAG TGATATTAGT TGGTGTCTCT CTGTTTATGT 840  
 10 CAAATATAAA TATTGTGGCG CATATTGGAG GATTCATTGG TGGTTTAITA ATAACCTTAA 900  
 TTGGCTAITA CTATAAAGTG AATCGTAATA TTTTTCGAT TTAATAATTG GTATGCTGT 960  
 15 CATATTTAWT GCACYTCAA TTAGATTTT ACATTAAAG AAGATAATAW TTATAA 1016

## (2) INFORMATION FOR SEQ ID NO: 467:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 406 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 467:

AACTTTAAAT TTAGACATCT TTAACACCTC TCTTAAACCA TGCCTATATC TCAAGATGAT 60  
 ATTTCAAATG AACAACTACT TTGCTTGAGA CCAITTAATGA ATGATCATAA ATATTTCTTT 120  
 30 CTATAAAATT AGCTTTCCAA TAACGTGTT GTTGCAATAT ATCATTACCA AGTACACCAT 180  
 TTTGGGAAGT ATGATTATCT TTATCTATAC TTAACCAAT TTGTTTAGTT TTAGCATGGC 240  
 TAAATTGTGT AAGACCCTTA CACGATAAAC GTATAGCGTC TGAATTCTCA TTTAACAATG 300  
 35 CGGCTGGGCA AACAAATGAC ACATTGTACT TCATGTTTGA ACTTCGTTAC AATCaTCGTG 360  
 kCaTTTGTat AAATAACAAT CCCTCGTAAT kGATTAAAGTA TATTAT 406

## (2) INFORMATION FOR SEQ ID NO: 468:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1378 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 468:

AGATAATACA CTTGAAGTTG GAATGGTTG TGACGGTTAT TTAATCGAA TTGAAAACTT 60  
 AACACCATCA AATTCTTCA ACTCAGCAAG TGAAGATACG ATTACTAAAA TTAATTTAAA 120

## EP 0 786 519 A2

|    |  |            |             |            |            |             |      |
|----|--|------------|-------------|------------|------------|-------------|------|
|    | AGGTACACGC                                 | TTAAACTAA  | GAGAAGCCAT  | CAATTATGAT | GAAATGGTTA | TTGTAGATAG  | 240  |
|    | TATGACGTAG                                 | TTCTTAATTA | TGCKAAAAGG  | GATTGATGAA | AAACTGAAGG | GCTTTTCATC  | 300  |
| 5  | AATCCCTTTT                                 | ATTTTAGGGG | AATTGAATAG  | ATAGTTTTAA | ACTATACGAA | TTATTAAATAT | 360  |
|    | TTGAGATTTA                                 | ATTGAAATAA | GTTTTAAAAA  | TTGGAGGAGA | TAGATTAAGC | GAAGTCATTT  | 420  |
|    | AAGGCTGAAG                                 | TTAAGTGTAT | TCACAAAAAn  | TAGCCACACT | CATATGACAT | CGGATGAGTG  | 480  |
| 10 | TGGCTTAAGG                                 | ATCTATGGGG | GGAGGAAnCC  | ATAGATGTTT | ACTTTGATAG | GCCAGATTAA  | 540  |
|    | ATATCAAAAT                                 | ATGCGATTAT | TTATAGCTTG  | ATGCAAAAGT | GGTATGCCTA | TTAAAAGTTA  | 600  |
|    | CTGCACATAG                                 | CTTTTAATAT | TCCGTTCAAA  | GGAAAGGGGC | ATACAATTGA | ACAATCTGTA  | 660  |
| 15 | ATAGTACTTT                                 | TAACCAGCTA | TGCTAAAAGT  | CTAGTAGGGA | GAACAGTTGT | CCAATCACAT  | 720  |
|    | AAGAACCCTC                                 | AACCTCGTTA | GTACGATTAA  | GAAAAGCTTT | TTAGTTAGTA | TGTAATACAA  | 780  |
| 20 | TTATTGACG                                  | CGCGTGAATC | TCTTTTATAA  | GAGTGTGTAG | GGAATGGCGT | TGTATAAATT  | 840  |
|    | GTATTAGAAG                                 | AACCTCTAAC | GCATCTCTGT  | GGTTAAAAGA | GATGAAGGGA | ACGACAGTTT  | 900  |
|    | aATTAAAACT                                 | GCATAAGAAC | TTCTAGCTTT  | TCTCTCTCGT | TCAAAGAGAA | GCAGCTGTTC  | 960  |
| 25 | GCAGTTTAAT                                 | CAAAACCACA | TAAAGCTTTT  | AACCTTACTC | TTTGATTTAA | AGAGTGATAA  | 1020 |
|    | ATGTTTACAG                                 | TTTAATTAAA | ACTGCATAAG  | AACCTCTAGC | TTTTCTCTTT | CGTTCAAGAG  | 1080 |
|    | AAGCAGCTGT                                 | TGCGAGTTTA | ATCAAAACCA  | CATAAAGCTT | TTAACTTTAC | TCTTTGATTT  | 1140 |
| 30 | AAAGAGTGAC                                 | AAATGTTTAC | AGTTTAAATTA | AAACTGCATA | AGAACTTCTA | GCTTTTCTCT  | 1200 |
|    | TTGTTCAAAA                                 | GAGAAGTTCT | AATACCACCA  | TATCGTGCGA | TCGGGAACGG | TATATATATT  | 1260 |
|    | AATAGGAGGG                                 | TAATATATAT | TTAACGCACG  | ATATGGGACT | ATTAGCCTTC | GACTTTGTTA  | 1320 |
| 35 | TGTTGATGTG                                 | TGGCCTAAAA | TATTGGAGAT  | ACCAATATTT | TAGGTTGCAT | CAACATCA    | 1378 |
|    | (2) INFORMATION FOR SEQ ID NO: 469:        |            |             |            |            |             |      |
|    | (i) SEQUENCE CHARACTERISTICS:              |            |             |            |            |             |      |
| 40 | (A) LENGTH: 4171 base pairs                |            |             |            |            |             |      |
|    | (B) TYPE: nucleic acid                     |            |             |            |            |             |      |
|    | (C) STRANDEDNESS: double                   |            |             |            |            |             |      |
|    | (D) TOPOLOGY: linear                       |            |             |            |            |             |      |
| 45 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 469: |            |             |            |            |             |      |
|    | TCCCAACCAA                                 | TAATCGTGGC | AAAAATACGG  | ATATTGGTAT | GGCTTACAA  | ATTGCAATA   | 60   |
| 50 | TCGTTTAATC                                 | ATACATCCCC | CCTAATCTAT  | TGCCCTATCC | TATTCaTAAG | CATAAAAAATG | 120  |
|    | AATAGAGGTT                                 | GGATACATAA | TTTGTAGATG  | TAAATCTTTC | TTACAATTTA | CATTTTTAAA  | 180  |

|    |  |      |
|----|--|------|
|    | TTATATTCTA TTCAATTAA TCTATGGATA CTGTGTCCTCC ACACGACAGC AAAAGTTATC  | 300  |
|    | ATACTTCTTT ACATCACTAA GTCAATATAA ATGATTTAAT CAGTATTAC ACTTTATTTG   | 360  |
| 5  | CTTAATACTG TCTAATTTT TTGTAACGTT CTTTCCAAAC TTGTATAAAA TCTGGCGGGA   | 420  |
|    | ATGGGCCCTT CTCTGTTCT ATCCATTGTT GAAGAATGTC CACGTTGGGT CTAAAAATAA   | 480  |
| 10 | TATCAATATC ATGGGATAA TTCATTGAT TCATATGTTG CTCATATCA TCTTCATCTA     | 540  |
|    | ATAAATGATA CTTTCCGTTT GGATATACTT TAATATCTAA ATCATAGTCT ATATATTTTA  | 600  |
|    | ATGCTCTCTC ATCACAACA AATGGTGATG ACAAAATGCA ATAGTAATAA ATTCCACTTT   | 660  |
| 15 | CTCTAAACAT GCAGATAACA TTAACCAAT ATTCTGAGTG AAAGTAACA ATTGCCGTTT    | 720  |
|    | CAOGTGTAT CCAAGTCTT COGTCACITT CAGTCACTAA CGTATGATCA TTTCCACCAA    | 780  |
|    | TGACAACATG ATCAGTACCC TTTAATATTG TTGTTTCAGA CCAAACGCGA TGAATCTTAC  | 840  |
| 20 | CATCATGTTT ATAACCTGTA ATTTTAATGT TTTCCCTTC TTAGGTATG GATTCTCTGA    | 900  |
|    | CCATACTCCA CACCACCTC TGTTAATTTA ACCATTATAA ATTATAGCAT ATTTTCAGAAA  | 960  |
|    | TAGTATTATA TAAATACATA TTTTACGAA ATAAGATTTT ACTACTTAAT AATTAAATCT   | 1020 |
| 25 | GGTAATATTG CTAAGTACTA CAACAGAGAT TTACATGTCC CATTTAAAGT ATATAAAATC  | 1080 |
|    | ATCACTTTTA TATATCAACA CTTTAACTTT TTGACATTGT TATCTATGA GATTTAAAGA   | 1140 |
|    | TATCATTTAT ACTTTTAA AATAATGTCA CTATGTTTC CGATAATAT ACCAATCATC      | 1200 |
| 30 | GAATGTTACC CATTATAAAA TTGATAAATs TTTGACATAG GTACAGGGAA TGTATATTGA  | 1260 |
|    | TCTOGATCAC TTAATCAAAA CCAATCATG TCATCTGGTA ATGTTTCAAT GTTAATTGCT   | 1320 |
|    | CCTGAAACGG CGTATACTTT AATCTTCCAT GTTAAATGAG TAAATIGATG CTTCAACTCA  | 1380 |
| 35 | AAAATAGGTG TTTCTACTGG TTGAATGTCA TGACCGATTT TTTAGTCAT TTTACGTCTA   | 1440 |
|    | GCATGCTCAC TTTCAACAT AGGAAATTGC CACATACCAT GCAATAATTT TTGCTACGC    | 1500 |
| 40 | TTTTGCAACA GATATTGACC TTGATTATTT CTAATTAAAA AGACGGATTG CTCGAATTACT | 1560 |
|    | TTTTACTTCA CATTTTIIAGA TTTAACAGGT AACTTTTCAA ATGTACCTTT ATCAAATGCC | 1620 |
|    | TCACAGTTTT CTGAACTGG ACAAATAAG CATATGGAT TTTTGGTGT ACAAATTAAAC     | 1680 |
| 45 | GCCCCTAATT CCATCATAGC TTGATTAAC GTTCCAGCTT CTGTAGTAAC ATACGGTAAT   | 1740 |
|    | AATCTGTGTT CGTACGATTT CCTCGTCGAT TGTAAATTAA TATCTCGATA GTCATATTC   | 1800 |
|    | AATCTAGACC ATACTCGAAA AACATTTCG TCTACGTTG CTAGTGGTAC ATTATATGCA    | 1860 |
| 50 | ATGCTCATTA CTGACGCTTG TGTGTATGGG CCAACACCTT TTAACGCTTT AAATTGATCA  | 1920 |
|    | GGATCTTTGG GAACTAAGCC TTCATATTTA TCATGAACCT CTTTAATCGC CGTATGAAAA  | 1980 |



|    |            |            |            |             |            |            |      |
|----|------------|------------|------------|-------------|------------|------------|------|
|    | TGTGCTATAC | CATTAACAAT | GTAATACATT | GGATTTAGCA  | TTAGGATGTG | ATTGATAAAT | 3900 |
| 5  | ACATGATTTG | GATTTGGTAT | GAATAATAAT | GGTAACAATA  | AGAAACACAA | TACACAAACC | 3960 |
|    | CCATAAATA  | TGATATTTAT | TTTTTCAGWT | AACAGTCGAA  | TAAAGACAAA | AGTAACGGAT | 4020 |
|    | ATTAATCCTA | CAAAAATAGT | TGCCATCACA | ATAAATAAGA  | ATAGCGCTAT | ATATGATGTT | 4080 |
|    | TCGAAGTTTA | CTGTGTTTAA | CAATGCACA  | ATCAATCGTCA | AAATGACATA | CATAATAAAA | 4140 |
| 10 | CTTAATAATG | ACATAATAAC | TACTGGGGTC | G           |            |            | 4171 |

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 9821 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

|    |             |             |            |             |             |             |      |
|----|-------------|-------------|------------|-------------|-------------|-------------|------|
|    | TGGTTGAAGT  | AGCAGTTAAT  | TCAAATCTG  | CAACAGTTTC  | AGCAGAATAG  | GGGCTTTCAA  | 60   |
| 25 | AATAACAAA   | GGAGAAATAT  | TTATGACTAA | AACTTTAAAG  | GTTTATAAAG  | GAGACGACGT  | 120  |
|    | CGTAGCTTCT  | GAACAAGGTG  | AAGGCAAAGT | GTCAGTAACT  | TTATCTAATT  | TAGAAGCGGA  | 180  |
|    | TACAACTTAT  | CCAAAAGGTA  | CTTACCAAGT | GGCATGGGAA  | AAAAATGGTA  | AAGAACTGAG  | 240  |
| 30 | TAAAGTTGAT  | GTAACCTCAAT | TCAAAGCCAA | TCCAATTCTA  | GTCTCAGGGC  | TATCATTTAC  | 300  |
|    | ACCAGAAACT  | AAATCAATTA  | TGTTAAATCA | CGATAGCCAT  | GTTGACCAAT  | ACATTGCACC  | 360  |
|    | AAGCAGACGA  | ACGAATAAAA  | TATTGAAATA | TACAAGTGAA  | CATTCAGAAAT | TGTGTACTGT  | 420  |
| 35 | AGATGAAAAA  | ACAGAGAGCA  | TTCAACGGTG | AGCTGAAGGT  | ACTTCAGTAA  | TCACTGCTAC  | 480  |
|    | GTCTACTGAT  | GGAAGCGATA  | AGTCAGGACA | AATTTTCAGT  | ACAGTAACAA  | ACGGATAGGG  | 540  |
|    | ATTTAAGGCG  | CAGTATATCT  | CGGCTTTTTT | TATTTGAATA  | AAAGGAGCTA  | ATACAATGAT  | 600  |
| 40 | TAAATTGAA   | ATTTAAGATC  | GTTAAACAGG | AAAAACGAG   | AGCTATACAA  | AAGAAGATGT  | 660  |
|    | AACAATGGGC  | GACAGAGAAA  | ATGCTATGAG | TATTTAGAAAT | TAGTAAATCA  | AGAGAAATAA  | 720  |
| 45 | AAAGAAGCAC  | CTAACGCAAC  | AAAAATGAGA | CAAAAAGAGC  | GACAGTTATT  | AGTAGATTTA  | 780  |
|    | TTTAAAGATG  | AAGGATTGCA  | TGAAGAAGAT | GTTCTGAACA  | AGATGAGTAC  | TAAAACCTTAT | 840  |
|    | ACAAAAGCCT  | TACAAGATAT  | ATTTTCGAGA | ATCAATGGTG  | AAGATGAAGA  | AGATTTCAGAA | 900  |
| 50 | ACTGAACCA   | AAGAGATGGG  | AAGACAGGAA | GACAACTGCT  | AATTAAGCAAT | ATTTTATCGA  | 960  |
|    | ACATTAGAGAA | AATACAACGT  | TTCTGTATGG | AGCATCATGT  | GTTGGACATTA | ACTGAAGTCA  | 1020 |

EP 0 786 519 A2

|    |  |      |
|----|--|------|
|    | AAGAAAAACA AAGTGAACAA AAGTCATTA CAGGTACGGA TTAAAGAAAA CTTTGTGGAA   | 1140 |
|    | GCTAGAAAGG AGGTTAATAT GAATGAAAAA GTAGAAGGCA TGACCTTGGG GCTGAAATTA  | 1200 |
| 5  | GACCATTAG GTGTCCAAGA AGGCATGAAA GGTTTAAAGC GACAATTAGG TGTGTGTAAT   | 1260 |
|    | AGTGAATGA AAGCTAATCT GTCAGCATTT GATAAGTCTG AAAAAATCAAT GGAAAAATAT  | 1320 |
|    | CAGGCGAGAA TTAAAGGGTT AAATGATAGG CTTAAAGTTC AAAAAAGAT GTATTCTCAA   | 1380 |
| 10 | GTAGAAGATG AGCTTAACAA AGTTAACGCT AATTACCAA AAGCTAAATC CAGTGTAAAA   | 1440 |
|    | GATGTTGAGA AAGCATATTT AAAGTTAGTA GAAGCCAATA AAAAGAGAAA ATTAGCTCTT  | 1500 |
|    | GATGAATCTA AAGAAGCCTT AAAATCATCG AATACAGAAC TTAATAAAGC TGAAAAATCAA | 1560 |
| 15 | TATAAACGTA CAAATCAACG TAAACAAGAT GCGTATCAA AACTTAAACA GTTGAGAGAT   | 1620 |
|    | GCAGAACAAA AGCTTAAGAA TAGTAACCAA GCTACTACTG CACAATAAA AAGAGCAAGT   | 1680 |
|    | GACGCTTACA GAAGCAGTCC GCTAAGCATA AAGCACTTGT TGAACAATAT AAACAAGAAG  | 1740 |
| 20 | GCAATCAAGT TCAAAAATA AAAGTGCAA ATGACAATCT TTCAAAATCA AATGATAAAA    | 1800 |
|    | TTGAAAGTTC TTACGCTAAA ACTAATACTA AATTAAAGCA AACAGAAAAA GAATTTAATG  | 1860 |
| 25 | ATTTAAACAA TACTATTAAG AATCATAGCG CTAATGTCGC AAAAGCTGAA ACAGCTGTTA  | 1920 |
|    | ATAAGAAAAA AGCTGCTTTA AATAATTTGG AGCGTTCAT AGATAAAGT TCATCCGAAA    | 1980 |
|    | TGAAGACTTT TAACAAAGAA CAAATGATAG CTCAAAGTCA TTTCGGTAAA CTTGCAAGTC  | 2040 |
| 30 | AAGCGGATGT CATGTCAAAG AAATTTAGTT CTATTGGAGA CAAATGACT TCCTTGGGAC   | 2100 |
|    | GTACAATGAC GATGGCGTGA TCTACACCGA TTACTTTAGG TTTAGGTGCA GCATTAAAAA  | 2160 |
|    | CGAGTGCAGA CTTTGAAGGG CAAATGTCTC GAGTTGGAGC GATTGCACAA GCAAGCAGTA  | 2220 |
| 35 | AAGACTTAAA AAGCATGTCT AATCAAGCGG TTGACTTAGG AGCTAAAAAA AGTAAAAGTG  | 2280 |
|    | CTAACGAAGT TGCTAAAGGT ATGGAAGAAT TGGCAGCTTT AGGCTTTAAT GCCAAACAAA  | 2340 |
|    | CAATGGAGGC TATGCCAGGT GTTATCAGCG CAGCAGAGC AAGTGGTGCA GAAATGGCTA   | 2400 |
| 40 | CAACTGCAAC TGTAATGGCT TCAGCGATTA ACTCTTTCGG TTTAAAGCA TCTGATGCAA   | 2460 |
|    | ATCATGTTCG TGATTACTT GCGAGATCAG CAAATGATAG TGCTGCAGAT ATTCAATATA   | 2520 |
|    | TGGGAGATGC ATTAATAATAT GCAGGTACTC CAGCAAAAGC ATTAGGAGTT TCAATAGAGG | 2580 |
|    | ACACTTCTCG AGCAATTGAA GTTTTATCTA ACTCAGGTTT AGAGGGGTCT CAAGCAGGTA  | 2640 |
|    | CTGCATTAAAG AGCTTCGTTT ATTAGGCTAG CTAATCCAAG TAAAGGTACA GCTAAGGAAA | 2700 |
| 50 | TGAAAAAATT AGTATTTCAT TTGTCGTATG CTAAGGTTGA GTTTGTGGA ATGGGCGAAT   | 2760 |
|    | TGATTAGACA GTTCCAAGAT AACATGAAAG GCATGACGAG AGAACAAAAA TTAGCAACAG  | 2820 |

55

EP 0 786 519 A2

CAGATAAAAT TAATAGCTAT AGCAAAATCAT TGAAGAACTC TAATGGTGAA AGTAAAAAAG 2940  
 CAGCTGATTT GATGAAAGAT AACCTCAAAG GTGCTCTGGA ACAATTAGGT GGCCTTTTGG 3000  
 5 AATCGTTAGC AATTGAAGTT GGTAAAGATT TAACGCCTAT GATTAGAGCA GGTGCGGAAG 3060  
 GATTAAACAA ATTAGTTGAT GGATTTACAC ATCTTCCTGG TTGGGTTAGA AAGGCTTCGG 3120  
 TAGGCTTAGC AATTTTGGT GCATCTATTG GTCTGCTGT TCTTGCTGGT GGCTTATTAA 3180  
 10 TACGTGAGT TGGAGCGCG GCTAAAGGCT ATGCATCATT AAATAGACGC ATTGCTGAAA 3240  
 ATACAAATCT TTCTAATACC AATTCAAAG CAATGAAATC TTTAGGCTTT CAAACATTAT 3300  
 TTCTGGGTTT TACAACAGGA AAAACGTCAA AAGGCTTTAA AGGATTAGCC GGAGCTATGT 3360  
 15 TGTTTAATTT AAAACCTATA AATGTTTTGA AAAATTCTGC AAAGCTAGCA ATTTTACCGT 3420  
 TCAAACTTTT GAAAACGGT TTAGGATTAG CCGCAAAATC CTTATTGCA GTAAGTGGAG 3480  
 20 GCGCAAGATT TGCTGGTGA GCCTTAAAGT TTTTAACAGG ACCTATAGGT GCTACAATAA 3540  
 CTGCTATTAC AATTGCATAT AAAGTTTTTA AAACCGCATA TGATCGTGTG GAATGGTTCA 3600  
 GAAACGGTAT TAACGGTTTA GGAGAAACTA TAAAGTTTTT TGGTGGCAAA ATTATTGGCG 3660  
 25 GTGCTGTTAG GAAGCTAGGA GAGTTTAAAA ATTATCTTGG AAGTATAGGC AAAAGCTTCA 3720  
 AAGAAAAGTT TTCAAAGGAT ATGAAAGATG GTTATAAATC TTTAGTGAC GATGACCTTC 3780  
 TGAAGTAGG AGTCAACAAG TTTAAAGGAT TTATGCAAC CATGGCACA GCTTCTAAAA 3840  
 30 AAGCATCTGA TACTGTAAAA GTGTTGGGGA AAGGTGTTTC AAAAGAAACA GAAAAAGCTT 3900  
 TAGAAAAATA CGTACACTAT TCTGAAGAGA ACAAAGAAAT CATGGA AAAA GTACGTTTAA 3960  
 ACTCGGGTCA AATAACAGAA GACAAAGCAA AAAAACTTTT GAAAAATTGA GCGGATTAT 4020  
 35 CTAATAACCT TATAGCTGAA ATAGAAAAAA GAAATAAAAA GGAACCTGAA AAAACTCAAG 4080  
 AACTTATTGA TAGTATAGT GCATTGATG AACAGAAAA GCAAAACATT TTAAGTAGAA 4140  
 40 CTAAGAAAAA AAATGACTTG CGAATTAAAA AAGAGCAAGA ACTCAATCAG AAAATCAAAG 4200  
 AATTGAAAGA AAAAGCTTTA AGTGATGGTC AGATTTCAGA AAATGAAAGA AAAGAAATTG 4260  
 AAAAGCTTGA AAATCAAAGA CGTGACATCA CTGTTAAAGA ATTGAGTAAG ACTGAAAAAG 4320  
 45 AGCAAGAGCG TATTTTAGTA AGAATGCAAA GAAACAGAAA TGCTTATTCA ATAGACGAAG 4380  
 CGAGCAAGC AATTAAAGAA GCAGAAAAAG CAAGAAAGC AAGAAAAA GAAGTGGATA 4440  
 AGCAGTATGA AGATGATGTC ATTGCTATAA AAAATAACGT CAACCTTTCT AAGTCTGAAA 4500  
 50 AAGATAAATT GTTAGCTATT GCTGATCAAA GACATAAGGA TGAAGTAAGA AAGGCAAAAT 4560  
 CTA AAAAAGA TGCTGTAGTA GACGTTGTTA AAAAGCAAAA TAAAGATATT GATAAAGAAA 4620

55



EP 0 786 519 A2

|    |  |      |
|----|--|------|
|    | GTGGTGGTC TAACTTTGA GAAGACCAA AGAAGAAAAG TGATAAATAC GCTAAAGAAC     | 4740 |
|    | AAGAAGAAAC AGCTCGTAGA AACAGAGAAA ATATAAAGAA ATGGTTTGA AATGCTTGGG   | 4800 |
| 5  | ACGGCGTAA AACTAAACT GGTGAAGCCT TTAGTAAAT GGGCAGAAAT GCTAATCATT     | 4860 |
|    | TTGGCGCGA AATGAAAAA ATGTGGAGTG GAATCAAAGG AATCCAAGC AAATTAAGTT     | 4920 |
|    | CAAGTTGGAG CTCAGCCAAA AGTTCTGTAG GATATCACAC TAAGGTCTATA GCTAATAGTA | 4980 |
| 10 | CTGGTAAATG GTTTGGAAAA GCTTGGCAAT CTGTTAAATC GACTACAGGA AGTATTTACA  | 5040 |
|    | ATCAAACTAA GCAAAAGTAT TCAGATGCCT CAGATAAAGC TTGGGCGCAT TCAAAATCTA  | 5100 |
|    | TTGGAGAGG CACATCAAAA TGGTTTAGCA ACGCATATAA AAGTGCAAGG GGTGGCTAA    | 5160 |
| 15 | TAGATATGGC TAATAAATCG CGCTCGAAAT GGGATAATAT TTCTAGTACA GCATGCTCGA  | 5220 |
|    | ATGCAAAATC CGTTTGGAAA GGAAACATCGA AATGGTTTAG TAATCATAAC AAATCTTTAA | 5280 |
| 20 | AAGGTTGGAC TGGGATATG TATTCAAGAG CCCACGATCG TTTTGATGCA ATTTCAAGTT   | 5340 |
|    | CGGCATGGTC TAACGCTAAA TCAGTATTTA ATGGTTTGA AAAATGGCTA TCAAAAACAT   | 5400 |
|    | ATGATTGGAT TAGAGATATT GGTAAAGACA TGGGAAGAGC TGGCGCTGAT TTAGGTAAAA  | 5460 |
| 25 | ATGTTGCTAA TAAAGCTATT GCGGGTTTGA ATAGCATGAT TGGCGGTATT AATAAAATAT  | 5520 |
|    | CTAAAGCCAT TACTGATAAA AATCTCATCA AGCCAATACC TACATTGTCT ACTGGTACTT  | 5580 |
|    | TAGCAGGAAA GGGGTAGCT ACCGATAATT CAGGAGCATT AACGCAACCG ACATTTGTCT   | 5640 |
| 30 | TATTAAATGA TAGAGGTTCT GGAAACGCC CAGGTGCTGG AGTTCAAGAA ATAATTCACA   | 5700 |
|    | GGGCTGACGG AACATTCCAT GCACCCCAAG GACGAGATGT GGTGTCTCCA CTAGGAGTTG  | 5760 |
|    | GAGATAGTGT AATAAATGCC AATGACACTC TGAAGTTACA GCGGATGGGT GTTTTGCCAA  | 5820 |
| 35 | AATTCCATGG TGGTACGAAA AAGAAAAAAT GGATGGAACA AGTTACTGAA AATCTTGCTA  | 5880 |
|    | AAAAAGCAGG GGACTTCGGT TCTAAAGCTA AAAACACAGC TCATAATATC AAAAAAGTG   | 5940 |
| 40 | CAGAAGAAAT GGTGAAGCG GCAGGCGATA AAATCAAAGA TGGTGATCT TGGTTAGCG     | 6000 |
|    | ATAAAATCGG CGATGTGTGG GATTATGTAC AACATCCAGG GAAACTAGTA AATAAAGTAA  | 6060 |
|    | TGTCAGGTTT AAATATTAAT TTTGGAGGCG GACTAACGCT ACAGTAAAAA TTGCTAAAGG  | 6120 |
| 45 | CGCGTACTCA TTGCTCAAAA AGAAATTAGT AGACAAAGTA AAATCGTGGT TTGAAGATT   | 6180 |
|    | TGGTGGCGGA GGCAGTGGAA GCTATCTATT TGACCATCCA ATTTGGCAAA GGTTTGGAG   | 6240 |
|    | TTACACAGGT GSACTTAACT TTAATGGCGG TCGTCACTAT GGTATCGACT TTGATATGCC  | 6300 |
| 50 | TACAGGAACG AACATTATG CTGTTAAAGG CGGTATAGCT GATAAAGTAT GGACTGATTA   | 6360 |
|    | CGGTGCGGTT AATCTATAC AAATTAAGAC CGGTGCTAAC GAATGGAAC TGTATATGCA    | 6420 |

55

EP 0 786 519 A2

|    |  |      |
|----|--|------|
|    | ATCAGGTGCT ACAGGTAAT TCCTTAGAGG AGCACACTTA CATTTCCAAT TGATGCAAGG   | 6540 |
|    | GTCGCATCCA GGGAAATGATA CAGCTAAAGA TCCAGAAAAA TGGTTGAAGT CACTTAAAGG | 6600 |
| 5  | TAGTGGCGTT CGAAGTGGTT CAGGTGTTAA TAAGGCTGCA TCTGCTTGGG CAGGCGATAT  | 6660 |
|    | ACGTCGTGCA GCAAAACGAA TGGGTGTTAA TGTACTTCG GCTGACGTAG GAAATATCAT   | 6720 |
| 10 | TAGCTTGATT CAACACGAAT CAGGAGGAAA TGCAGGTATA ACTCAATCTA GTTCGCTTAG  | 6780 |
|    | AGACATCAAC GTTTTACAGG GCAATCCAGC AAAAGGATTG CTTCAATATA TCCACAAAC   | 6840 |
|    | ATTTAGACAT TATGCTGTTA GAGGTACAAA CAATATATAT AGTGGTTACG ATCAGTTATT  | 6900 |
| 15 | AGGTTCTTTT AACACAGAT ATTGGCGCTC ACAGTTTAA CCAAGAGGTG GTTGGTCTCC    | 6960 |
|    | AAGTGGTCCA AGAAGATATG CGAATGGTGG TTTGATTACA AAGCATCAAC TTGCTGAAGT  | 7020 |
|    | GGGTGAAGGA GATAACAGG AGATGGTTAT CCTTTTAACT AGACGTAAAC GAGCAATTCA   | 7080 |
| 20 | ATTAAGTAA CAGGTTATGC GCATCATCGG TATGGATGCG AAGCCAAATA ACATCACTGT   | 7140 |
|    | AAATATGAT ACTTCTACAG TTGAAAAATT GTTGAAACAA ATTGTTATGT TAAGTGATAA   | 7200 |
|    | AGGAATAAA TTAACAGATG CGTTGATTCA AACTGTTTCT TCTCAGGATA ATAACTTAGG   | 7260 |
| 25 | TTCTAATGAT GCAATTAGAG GTTTAGAAAA AATATTGTCA AAACAAAGTG GGCATAGAGC  | 7320 |
|    | AAATGCAAAAT AATTATATGG GAGGTTTGAC TAATTAATGC AATCTTTTGT AAAAATCATA | 7380 |
| 30 | GATGGTTACA AGGAAGAAGT AATAACAGAT TTAATCAGC TTATATTTTT AGATGCAAGG   | 7440 |
|    | GCTGAAAGTC CAAACACCAA TGATAACAGT GTAACATTA ACGGAGTAGA TGGTATTTTA   | 7500 |
|    | CCGGCGCAAC TTAGTTTTCG GCCTTTTTC AATTGATTAA GGTGTTGGCTA TGATGGTATA  | 7560 |
| 35 | GATGTTATAG ATTTAAATTT ATTTGAGCAT TGGTTTAGAT CTGTGTTTAA TCGCAGACAT  | 7620 |
|    | CCTTATTATG TTATTACTTC TCAAATGCCT GGTGTTAAAT ATGCAGTGAA TACAGCTAAT  | 7680 |
|    | GTTACATCTA ATTTAAAAGA TGGTTCCTCA ACTGAAATTG AAGTAAGTTT AAATGTTTAT  | 7740 |
| 40 | AAAGGGTATT CTGAATCAGT TAATGGACC GATAGCGAGT TCTTATTGCA CTCTAATTGG   | 7800 |
|    | ATGTTTGAAA ATGGAATTCC TCTTGATTTC ACACCTAAAT ATACTCATAC ATCAAAATCAA | 7860 |
|    | TTTACTATTT GGAACGGTTC TACTGATACG ATAAATCCAC GATTCAAGCA CGATTTGAAA  | 7920 |
| 45 | ATATTAAATTA ATTTAAATGC GAGTGGAGGA TTTGAACTGG TTAATTATAC AACAGGTGAT | 7980 |
|    | ATTTTAAAGT ACAACAAAG TATAGATAAA AACACTGATT TTGTTTTAGA TGGTGTGTAT   | 8040 |
| 50 | GCATATCGAG ATATAACAG AGTGGGAATT GATACAAATA GAGGCATTAT AACATTAGCG   | 8100 |
|    | CCAGGTAAAA ATGAATTTAA GATTAAAGGA GACGTCAGTG ATATTAAAC TACATTAAAG   | 8160 |
| 55 | TTTCCTTTTA TTATAGGTA GGTGATTTAA TGGATTATCA TGATCATTTA TCAGTAATGG   | 8220 |

|    |             |             |             |             |            |             |      |
|----|-------------|-------------|-------------|-------------|------------|-------------|------|
|    | ATTATGAAC   | GAATGAAGCT  | AGGTACATCA  | CCTTTACAGT  | TTATAGAAGT | ACTCATAATA  | 8340 |
|    | GTTTGTGTTT  | TGATTATATG  | ATTGTGTGAA  | ACTTCATAAT  | TTATCATGGT | GAAGAAATACA | 8400 |
| 5  | CAATTAAAGCA | GACAGCGCCA  | AAGGTGAAG   | GTGATAAAGT  | TTTTATTGAA | GTTACGCGAT  | 8460 |
|    | ATCACATAAT  | GTATGAATTT  | CAAAATCACT  | CAGTGGAAATC | AAATAAGCTT | GATGACGACA  | 8520 |
| 10 | GTAGCGAAAC  | TGGTAAACG   | CCGAATACT   | CTTTAGATGA  | GTACTTAAGA | TATGGATTTG  | 8580 |
|    | CAAAATCAAAA | AAGCTCAGTC  | AAGATGACCT  | ATAAAATAAT  | TGGAGATTTT | AAAAAGAAAA  | 8640 |
|    | TACCAATTGA  | TGAATTAGGT  | AATAAAAAATG | GCTTAGAATA  | TTGTAAGAG  | GCAGTAGATT  | 8700 |
| 15 | TGTTTGGTTG  | TATTTATTTT  | CCAAATGATA  | CGGAGATGAT  | TTTTTATCT  | CCTGAAACAT  | 8760 |
|    | TCATCAAAAG  | AAGCGAAAAA  | GTAATAAAGT  | ATCAATATA   | TACTGATCT  | GTGTCTGCTA  | 8820 |
|    | CTGTCAAGTAC | GTTGGATATA  | AGCAACAGTA  | TAAAGTTTT   | TGGGAAAAAG | TACACAGCCG  | 8880 |
| 20 | AGGAAAGAG   | AAATTATAAT  | CCTATTAGAA  | CAACTGCATC  | TAAATATTCA | AATGGTTTTA  | 8940 |
|    | TAAAGAGAGG  | TACTTATCGT  | ACCGCAACAA  | TGGGGTCTAA  | AGCTACTATT | AACCTTGATT  | 9000 |
|    | GCAAGTATGG  | TAATGAAACA  | GTTAGATTTA  | CAATAAAAA   | GGGCTCTCaA | GGTGGAAATAT | 9060 |
| 25 | ATAAGTTGAT  | TTTAGACGGC  | AAGCAaATT   | AGCAaATTTC  | TGTTTGTGCT | AAGTCGGTTC  | 9120 |
|    | AGTCTGmAC   | AATGATTTTA  | ATaAAAAATA  | TTGATAAAGG  | CAAGCACGTT | TTAGAAATGA  | 9180 |
|    | TATTTTTTGG  | AGaGfCCCC   | AAAAATAGAA  | TTGATATATC  | TTCAAAATAA | AAAGCTAAGC  | 9240 |
| 30 | CTGTATGTA   | TGTTGGAACT  | GA AAAATCAA | CAGTCTTAAA  | TTTAATTGCT | GACAACCTAG  | 9300 |
|    | GTGCAATCA   | ATACAAAGCA  | ATTGTTGaCT  | ACGTGcAGA   | TAGTGCAAG  | CAGTTTGGGA  | 9360 |
| 35 | TTCGATATGC  | TAATACGCAA  | ACAAATTGAAG | ATATCGAAAC  | ACAGGATAAG | CTGTTAGAAT  | 9420 |
|    | TTGCAAAAAA  | GCAAAATAAT  | GATACTCCTA  | AGACTGAATT  | AGATGTTAAT | TATATAGGTT  | 9480 |
|    | ATGAAAAAAT  | AGAGCCAAGA  | GATAGCGTAT  | TCTTTGTTCA  | TGAATTAAAG | GGATATAACA  | 9540 |
| 40 | CTGAATTAAAG | GTTTGTAAA   | CTTGATAGGT  | CACATCCATT  | TGTAAACGCA | ATAGATGAAG  | 9600 |
|    | TGTCCTTTAC  | CAATGAAATA  | AAGATATAG   | TACAAATTCA  | ACAAGCGCTT | AACAGACGAG  | 9660 |
|    | TTTCTGCACA  | AGATATATAGA | TATAACTATC  | AGCAAAATCG  | TATAAATCAT | TTATACACTA  | 9720 |
| 45 | GTACTTTGAA  | TTCTCCTTTG  | GAGACAATGG  | ATATAGGGAG  | TGTAATTAA  | TAATGGCAAC  | 9780 |
|    | AGAAGAAGTT  | AAAAATCAAG  | CGCTACTTGA  | AAACGATAAA  | C          |             | 9821 |

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1017 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 471:

35 (2) INFORMATION FOR SEQ ID NO: 472:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6806 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEO ID NO: 472:

|    |  |     |
|----|--|-----|
| 45 | TAGGATTGAT TAATCCTTTC TTCAAAATGA TGAATGTGTA ATGTAAATA TATAATTCCA     | 60  |
|    | GACTCACTGA TATTAAACATC AAATTGTTTT TGTATCATAG TTA AAAATTTT ATATGCTGTG | 120 |
| 50 | TTATAGCAAA TCGGATAGTG ATTTTAAATC ATAGACACAA AATCATCTTG TGCATGTATA    | 180 |
|    | TATTCCTTTC TTCTTAAATG GCGAATTAAA AATTGTACGT GCCTTATAAA ACGTTGGTAT    | 240 |

EP 0 786 519 A2

|    |   |      |
|----|---|------|
|    | TTAATAACAT TATThGATCA AGGTCATCTC ATGCATAGAT AAATCTTCTG TATTaGATGC   | 360  |
|    | AATATGTAAT GCAATAAACT CTATCTCATC TTCAGGAAAA TGTaCATCTA ATGCTGCATT   | 420  |
| 5  | TAAC TGATCA ATCACCTGTT TAGCAATATG ATATGCATCA CTATATAACT GCATAGTTTC  | 480  |
|    | CATACAAAT GGATTGCTAA TAAC TTGATT TTGTTTAAAG CGTTTATAAG CAAATATAAT   | 540  |
|    | ATGATCCGTT AATGAAACTA CAAGTTGTTT TGAATCAACA TTCATCGCAG TATTAGAAAT   | 600  |
| 10 | AAAATTCAAC GAATCAATAA TTACTTGTA TACATTATCA TCAGCGATTT CAACTAAACT    | 660  |
|    | TTTATAATGT GCTTTTGTG GCTCACITTC TAATTATATA ATTTTCTCTA TTGTAATAGT    | 720  |
|    | TTGTCGTTT AACGCCATTC CCTCTTTTTT GTTAAAACCA ATACCTTTAC CGATTAAAAAT   | 780  |
| 15 | AAC TTCTGA TCATTATTAG TACATACTAC GACATTGTTG TTCAATGTTT TAGTAACAAT   | 840  |
|    | ATATTCTCCC ATTATCATCA CCTATTTTTT TATTATTAAG ATTATATATC GGAAATGTCT   | 900  |
| 20 | AAC TTGAAAG TATAATAATT TAAATACTTG TGTTCTATTA CACGCCATTC CTATATGATA  | 960  |
|    | TATCTTAATT TAATTTCGAT GTCTCTCAA GTGGAATAAC TATAATAAAA AAATCTGACT    | 1020 |
|    | CCCATATTTA CAAATAATTC TATTTATTTA TATATTATCT GAATTAATAC TCAATACAAA   | 1080 |
| 25 | CTAAAACGTA CTATTAAATT GTGCAAGCT AAAACAAATT TATATTATCT TATCCAACAA    | 1140 |
|    | TATGTCTTAT CAATGGTATA GTCTTTGCAC ACCAATGGAG GAAATAAATC TCAACCTTAC   | 1200 |
|    | TATATTAATA TATAATCAA TCTTAGATTA ACTAGTGTA TGTATACAGAT GATAATTGAG    | 1260 |
| 30 | TACAAATTTA AAACCCGAG ATTTTCGCTT TAATTTGAAA ACCTCAGGCT TTATTGGATT    | 1320 |
|    | TTTATATAAT GAATCGTTAC ATTAATAATA TTTATTATC AGAGTTCTTA TATTGTTAG     | 1380 |
|    | CGCCCCAGC ACTAATTCCA AATAAGTTAA TTCTAAGTT TTCAGGTTTA AAGACAGGCT     | 1440 |
| 35 | TCTTGCCITC TTTTTTCGTC TTTTGATAAT CTTTCATCAA TGCAAAAGCT ACATTGGACA   | 1500 |
|    | GTCTTATAAT GGAAATAATG TTTACAATTG CCATTAAGCC CATAAATAAG TCTGCCGTAT   | 1560 |
| 40 | TCCATACTGT TTCTGTTTTT ACAACTGCAC CGACAAAGAC AAGTACTACA ACAAGACATC   | 1620 |
|    | TAAAGATAAA TAATATTACA CGGTTTGTTG ATAAAAATTC AATATTAGAT TGACCGTAAT   | 1680 |
|    | AGTAATTACC TACAACAGAT GAAAATGCAA ACAGTGTAA tGCTATTGTT AAGAAAAATC    | 1740 |
| 45 | CTCCAGCAGA ACCTAAATGC TCATTAAATG CTGATTGAGT AACTGCAACA CCTTGAGGTG   | 1800 |
|    | CGTTATCACC AAATTTTCAGT CCTGAATATA GTAAAAATCAT GATTGCAGIT GCTGTACAAA | 1860 |
|    | CCAACATTGT ATCAAGAAG ACACCTAATG ATTGGATTAA ACCTTGCTTA ACAGGGTGTG    | 1920 |
| 50 | GTACGCGAGC AGTTGCCGCT GCATTCCGCG CAGAACCCAT ACCAGCTTCG TTAGAGAATA   | 1980 |
|    | AACCACGTTT GATACCTTGA AGAACCGCAG CACCTACAGC GCCACCAGTT ACTTGTTCSA   | 2040 |

55

|    |  |      |
|----|--|------|
|    | GCAATATTAC TAAAACCATA CCAATGTAAA TGATAGCCAT AATCGGTACA ATTAACGAAG  | 2160 |
|    | ATAACGTAGC AATACTACGT ACACCACCAA ATATAATAAT AGCTGTTACG ATTGCTAAAA  | 2220 |
| 5  | TAATACCTGT GATTACTGGA CTAATATTAT ATTGCGTATT TAAGCACTCC GCAATTGTAT  | 2280 |
|    | TAGATTGCAC TGTGTTAAAT ACAAATGCAA ATGTAATTGT AATTAAATC GCAAAATGAA   | 2340 |
|    | TACCTAGCCA TTTTGTATT AAACCTTTAG TAATATAGTA AGCTGGACCA CCACGGAATC   | 2400 |
| 10 | CACCATCTTT ATCATGTACT TTATAAACCT GAGCCAAAGT CGCTTCTATA AATGCACTCG  | 2460 |
|    | CTGCACCTAT AAATGCAATA ACCCACATCC AAAAATCTGC ACCTGGACCG CCTAAACAAA  | 2520 |
|    | TCGCAGTCGC AACACCAGCA ATATTACCAG TACCAACTCT CGAACCGACA CTAATCGCAA  | 2580 |
| 15 | ATGCTTGGAA TGGCGAAATA CCCTTCTTAC CATCTTCTAA AGTTTCTGGA GTTCTACTA   | 2640 |
|    | AAGCTCTAAA CATTTCAAGT AACATTGCTA ATTGAACGAA TTTAGAATA ATCGTAAAGA   | 2700 |
|    | AGAATCCAGC TGTCATAAT AGACCAATTA AATATTGAGA CCATATTAAA TCGGTACCAA   | 2760 |
| 20 | CATGGACAAA TTCTTTAAAC CATCCAGGTA TTAACATATC GAAATCTTTC AAATAAAACC  | 2820 |
|    | CCTCGCATCC TCTACATGAA TCATGTACCT TCTATAAAAT TAGACCGAAT TGAACCTTCA  | 2880 |
| 25 | GTAATATATG AGATACATCA TCATTTCTTA TACAATACAA GAGATTATATA TTAGTTTGGT | 2940 |
|    | CMAAGTATAT CGCTAATTTA ACGATAAGTA CTGGTCAGC ATTTAATATA AATCCCTTGA   | 3000 |
|    | ATTIAGTCAA AATTTAACAT TACTGTATTT TATCATTTAA TTTCTGTATT GCATATAGTT  | 3060 |
| 30 | TTTAGCTAAT ATACATGTCT ATTACTTCAC CAAAATCATC TGATCTACA ATGAATGAGC   | 3120 |
|    | CATTTGTATA TTGTCAGAT TTATGAATAT CATTAAATTA ACCATGTTCT TCATTTGATT   | 3180 |
|    | TTGAATATAA TGTATATTGA CTATGTTTAC CTGTCACTAC ATGTGCACTC ACAAATGAT   | 3240 |
| 35 | GTGGATTTTT CTTTAATTCT TTTAATAAAG TTATTCCACG TTGTGCTCTT TTAGCAACTT  | 3300 |
|    | GTAAGATTTT AAAACTAATA CGTTTTAACG AGCCGCGTTG TGTGGCCATC AATATAGTAT  | 3360 |
|    | CATTTTCAGA AACACCTTCT GTCATACCAA CGAAATCTTC AGCTTTAAGA TTTATTGATT  | 3420 |
| 40 | TAACACCAGC TGCCCTTAAT CCGGTATCTG ATAGTTCACT TGTATTATAC GTTAATGACA  | 3480 |
|    | TACCTTTATT AGTAATGACG GTAATTAAAT GATCTTTTTC AAAGCGCATA ACATAATCA   | 3540 |
| 45 | AATCATCATT TTCTTTAACT TTAGTAGCAA TTAAGGTTT ATTAAGAACGC GTTGTTTTAA  | 3600 |
|    | ATAGAGGCAC TGTACTTTTC TTAATCATGC CATTTTGAGT CGCAAAAACA TAAATGAT    | 3660 |
|    | CTGTATTAAA GTCCCTTTTC TTAAGAGCAT TAATAACAC TTCACTTCTC TCGATAGGAA   | 3720 |
| 50 | CTATTGTGTA TCAATGTTGT CCCAATTCTT TCCAACGAAT ATCTGCTAAT TTATGAACCG  | 3780 |
|    | GTATAAATAG ATAACGACCT TTATTGTGTA ATACTAGTAC GGTATCTTGC GTATTTACTT  | 3840 |

EP 0 786 519 A2

|    |  |      |
|----|--|------|
|    | TAAAGCTACG AATAGAAGTA CGTTTAATAT ATCCATGACG TGTCACTT AAAATACTy     | 3960 |
|    | CTTCACTAGG CACCATAACT TCTTTGTCAA TTTTAATTTC TTCAATTCTT GCTTCAATTA  | 4020 |
| 5  | AAGACAGTCG TTCAGATTG AATTTCCTTT TAATTCATT CAATCTCTCT TTTATGACAT    | 4080 |
|    | TCAATATATC ATCATGGTTA TCAAGAATAT GACGTAATTG TTTGATTAAAT GCTTCAAGTT | 4140 |
|    | CTTTATGTC ACCTTCAAGC GCAACTATGT CAGTATTGTG TAAACGATAT AACTGTAAAC   | 4200 |
| 10 | TTACAATTGC TTCAGCCTGT TCTTCTGTGA ACTCGTATAC TTCGATAAGG TTTTCTTTAG  | 4260 |
|    | CGTCACGCTT GTTTTTAGAG CWAOGAATCA ATTCGATTAC TTTATCTAAA ATTGACAACG  | 4320 |
|    | CTTTAATCAA ACCTTCAACG ATATGCATAC GTTTTCTGCG ATTATCTAAT TCAAACCTCG  | 4380 |
| 15 | TTCTATTTCG AACCACTCA ATTTGGTGAT TCAAATAACT ATCTATAATT TGACGAATAC   | 4440 |
|    | CCATCAATTI TGGACGACCA TCACTAATAG CGACCATGTT GAAATTATAT GAAACTGTGA  | 4500 |
| 20 | AATCAGAGTT TTTATAAAGA TAATTTTGA TTGATTCAT GTTCACATCT TTTTTCAAAT    | 4560 |
|    | CAATTGCTAT TCGTAAACCA GTTCTATCAG TTTTCATCAG TACTTCAACG ATACCATCGA  | 4620 |
|    | CTTTTTTGTC AGCACGTAAAT TCATCGATAC GTTTTACTAA GCTACTTTTG TTCATCTCAT | 4680 |
| 25 | ATGGAAATTC AGTAATAATT AACTGTTTAC GTCCATTGCG TAAAGTTTCT TCTTCAACTT  | 4740 |
|    | TAGAACGAAC TATAATTCTA CCTTTACCTG ATTCATAAGC TTTTAAATA CCATCAATAC   | 4800 |
|    | CTTGAATTAAT ACCACCAATT GGAAAATCAG GACCTTTAAT ATATTTCATT AATTGATTGA | 4860 |
| 30 | CTGTAATATC CGGATTATCA ATATATTTAA GTGTGCTTG AATCACTTCA GCTAAATTAAT  | 4920 |
|    | GTGGTGGTAT ATCTGTCGCG TAACCTGCAG ATATACCTGT AGAACCATTC ACTAGTAAGT  | 4980 |
|    | TAGGAAATCT TGATGGCAAT ACCATTGGTT CGAGTGTCTG ATCATCATAG TTTGGAATGA  | 5040 |
| 35 | AAGAAACTGT CTCCTTATTA ATATCACGTA ATAACTCTTC AGCTAGTAAG CTAACTTAG   | 5100 |
|    | CTTCAGTATA ACGCATTGCC GCTGGCGGAT CATTATCGAT ACTACCATTA TTACCATGCA  | 5160 |
| 40 | TTTCTATTAA GACATGTCGT AACTTCCAGT CTTGACTTAA ACGGACCATT GCTTCGTACA  | 5220 |
|    | CTGAGGAGTC TCCATGTGGA TGATATTGAC CAATAACATC ACCGACTGTT TTCGCACTT   | 5280 |
|    | TACGGAAATT TTTATCGTGT GTATTACCAC TTGAATACAT TGCATATAAA ATACGACGTT  | 5340 |
| 45 | GTACTGGTTC TAAACCATCA CGAATCATCT GCAATGCACG CTCTTGAATA ATATATTAC   | 5400 |
|    | TATATCTTCC AAACGATCA CCTAAAACAT CTTCAAGTGA TAAACTCTGA ATTATTTCAC   | 5460 |
|    | TCATAGATTG TCTCTCTCAT CAAATTGATC ATTTTCAAGC ACTTGACTT CAGAATTATC   | 5520 |
| 50 | TAAAATACTT TGGTCTCTT GCATACCAA CTCAACATGC TTTTCAATCC ATTCACGTCT    | 5580 |
|    | AGGTTGTAAT TTGTCACCCA TTAATGTTGT TACACGTTTA GATGAACGCA CTTCATCTTC  | 5640 |
| 55 |  |      |

AGGGTTTCATT TCACCCAAAC CTITGTAACG TTGTAACGTG AaGCCTTTAC CAAGTTCCTT 5760  
 TTGCAATTIA TTAAGCTCTT CGTCTGTCCA AGCGTATTCA ACTCGCTTTG TTTTGCCCTT 5820  
 5 ACCTTTTTC C AATTTATAAA GTGGAGGTAA AGCAATAAAT ACACGACCTG CTTGAACAAG 5880  
 CGGTTTCATA TATTTGAAGA AGAATGTAA CAATAGCACT TGAATATGCG CACCATCAGT 5940  
 ATCAGCATCA GTCATATAA TTACACGATT ATAATTACTA TCTTCAATTT TAAAGTCAGT 6000  
 10 ACCAAGCGCT GCCCCGATTG TGTGGATAAT TGTATTAAAT TCTTCAATTT TAAAAATATC 6060  
 TTCTAGACGT GCTTTCTCTG TATTAAATTAC CTTACCACGT AATGGTAATA TCGCTTGGAA 6120  
 TTTGCGGTCT CGTcCAAGTT TTGCTGAACC TCCCGCAGAA TCACCTTCGA CTAATACAA 6180  
 15 TTCATTTTTT TCAGTGTITT TACTTTGTGC AGGTGTTAAT TTACCAGATA GCAAAGTGTC 6240  
 TTTACGCTTG TTTTCTTAC CTGAACGAGC ATCTTCACGA GCTTTACGTG CAGCTTCCCT 6300  
 TGCTGTGTG GCTTAAATCG CTTTTTTCAC AAGTGATTGA GACAATTGTC CTTTTTCTTC 6360  
 20 TAAATAGAAAT GGCAATTTGT CTGCAACAAC TGAATCAACA GCACCTCTAG CTTCAGAAAGT 6420  
 ACCCAATTTA GATTTCGITT GTCCTTCAA TTGCAATAAT TCTTCTGGAA TACGAACAGA 6480  
 CACAACAGCT GTTAAACCTT CACGAATATC ATTACCATCT AAGTTTTTAT CTTTTGTTTT 6540  
 AAGTTCATTA ATACGACGTG CATAATCATT AAATACACGT GTCATTGCTG TTTTAAACC 6600  
 AACTTCATGT GTACCACCAT CTTTAGTAGG TACATTATTT ACAAAACCTA AAATACTTTC 6660  
 30 TGAATATTGA ACATTATATT GGAAGCTAC GTCTACCTCT ATACCATTG CTTCACTGA 6720  
 AAATGTAGCC ACGTCATGCA AAACCTCTTT TCCTTCATTG ACATAACTAA CAAACTCTTT 6780  
 GATTCCTTCT TATAATGGTA TGTCTT 6806

## (2) INFORMATION FOR SEQ ID NO: 473:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1716 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 473:

GGGGCAnAAA TTCCAATACA CTCATTACCA AATATATACA CCACTTCCTG CTACAAGTnn 60  
 TTTTACTTGA TCTTGGTCTT TTCCGCAGAA AGAGCAATTC CAAATTTCTC TCATCTTCCA 120  
 50 TTTGAATTTAA ACATTCTTTT TACACCCCTA TCGCTTAAAG ACTATACTAG ATTGGATGTT 180  
 ACAATGCAAC ATATTAAACAT ACAAACTTTT TGCTTAAAGA ATAGTAGCAG ATACATAAGC 240



EP 0 786 519 A2

|    |   |      |
|----|---|------|
|    | CCCTCAACGA ACTTTGCGTT ATCTCTTAAT AAATCGATAA CTTTTGGAT ACGAACATCA  | 360  |
|    | TTTTTAATGA TATCAGTATT ACCTAAAGTA TTTTGTATAT CTTCACCTGA GATATTAAAT | 420  |
| 5  | TGTTTACTCA TTTTTCCTAA TTCTTTATCG ATATCTTCAT CAGTAGCTTC GATTTTTTCA | 480  |
|    | GCTTCAGCGA TCGCAGTTAA AGTTAAGTTA GTTTTAACAC GTTGTCTCTG ATCGTCTTTC | 540  |
|    | ATTTCGCTCT TTAATTGAGT TTCATCTTGA CCTGAGATTG GGAAGTACGT TTGTAATCTC | 600  |
| 10 | AAACCTTGTT GTTGAATTCT TTGTGCAAT TCAGACACCA TACGATCTAA TTCAGTATTA  | 660  |
|    | ACCAATGCTT CAGGAATATC GATTGTTGTA TTATCAGTAG CTTTGTGAAT CGCTTCTTCT | 720  |
|    | TTTTCAACAT TTTCAGCATC TGTAAGCTTT TTGTTCAGCT AACGTTTACG TAAGTTTTCT | 780  |
| 15 | TGTACTCGT CTACTGTATT TGCTTCTGCA TCTAATTCAT TAGCAATTC ATCTGTTAAT   | 840  |
|    | TCTGGGACTT CTTTAAATTT AATTTCGTTA ACTTTTGTGT TGAAAGTTGC TTCTTTACCG | 900  |
| 20 | GCTAATCTTT CAGCATGGTA TTCTTCGGG AATGTTACGA CAACATCTTT TTCTTCGTCA  | 960  |
|    | ACTTTCATAC CTCTTAATTG CTCTTCGAAA CCAGGTATGA ATGAACCTGA ACCGATTCTT | 1020 |
|    | AAATCGTAAC CTTCAGCTTG TCCACCTTCG AATTCTTCTC CGTCAACTGA ACCACTAAAG | 1080 |
| 25 | TGCAATGTAA CTGTGTCGCC ATTTTCAACA ACACCATCTT CTTTAAAGCA CATTTCAGCT | 1140 |
|    | AAATGTCCTA AGCTGTGGTC AATCGCTTCT TGTAACCTAT CATCAGATAA TTCAGTTTCT | 1200 |
|    | TGTTTTTCAA TTTCAGACC TTTATAGTCT CTAATTAA CTTCGGCTC AACTGTAAT      | 1260 |
| 30 | GTGCTTCAA AATGAAATC TTTACCTTTT TCAATTGAG TAACACTTAC TTCTGGTTGT    | 1320 |
|    | GCAACTGGTT TAATATCAGT TTGTCGAAT GCTTCACCAT AAGCATCTGG TAATAAAATG  | 1380 |
|    | TGATAGCAT CTGATATATA TGCTTCTACA CCAAAGCGIT GTTCAAAAAT TGGACGTGGC  | 1440 |
| 35 | ACTTTACCTT TACGGAATCC AGGTACGTTA ATTTGTTTAA CCACCTTTTT GAATGCTTGA | 1500 |
|    | TCTAACGCTT TGTTTACTTT TTCTGCAGGA ACAGTAACAG TTAATAAAAC TTCGTTACCT | 1560 |
| 40 | TCCTTTTTTT CCCAAGTGC TGTCATGTAT ATATACCTCC ATGATTAACT AATTTATTTT  | 1620 |
|    | TTCAACTTCC CTATTATATC ATACGTCAT TCCCTATACA AACATTGAAA TCACAACGTT  | 1680 |
|    | TATATATTGG TAAATCAACT TTTTTCGTCA AAACCTA                          | 1716 |

(2) INFORMATION FOR SEQ ID NO: 474:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 795 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

TGACCAAGTA CGTTTCGAAG TTGCCATTAA AGCATTAAAC CCATCATTGA AAGCATTTCG 60  
 ACCTGTACGT GAGTGGGCAT GGAGTCGTGA AGAAGAAATC GATTATGCAA TTAACATAA 120  
 5 TATCCCTGTA TCAATCAACC ATGATTACCC TTATTCCTATC GATCAAAATC TATGGGGCAG 180  
 AGCGAATGAA TGTGGTATTT TAGAAGATCC TTATGCTGCG CCACCAGAGG ATGGGTTTGA 240  
 TCTAACAAAT GCCTTAGAAG AAACACCAGA TACTGCTGAT GAAATCATTT TAAAGTTTGA 300  
 10 TAAAGGCATC CCAGTTCAAA TTGATGGCAA AACATATGAA TTAGACGATT TAAATTTAAC 360  
 GTTGAATGCA TTAGCTGGTA AGCATGGTAT CGGAAGAATT GACCATGTAG AAAATAGACT 420  
 TGTAGGTATC AAATCAAGAG AAATTTATGA GGCACCTGCT GCAGAAGTTA TTTTAAAGC 480  
 15 GCATAAAGCA TTGAAACGA TTACGTTAAC GAAAGATGTC GCACACTTTA AACCAATCAT 540  
 TGAGAAGCAA TTTGCTGAAC AACTATACAA TGGACTTTGG TTCTCACCTT TAACTGATAG 600  
 CTTGAAATTA TTTATTGATA GTACTCAGCA ATACGTTAAG GGTGATGTCA GAATTAATTT 660  
 20 ATTCAAAGGT AATGCCATCG TGAATGGTAG AAAATCACCT TACACATTAT ATGATGAAAA 720  
 ATTAGCAACT TATACAAAAG AAGATGCATT TAAATCAAGC GCTGCTGTGG GCTTTATCGA 780  
 25 TATCTATGGT TTACC 795

## (2) INFORMATION FOR SEQ ID NO: 475:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 887 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 475:

CGATTGAAAG AAGACGGTTC AGTTGAAAAG TTTCCAAAGC CAGTAATTAG CCAACAACCA 60  
 GAAGGATATA CGATCATTT TAGAGATCCT AAAGTTTTTA AATATGATG GAATATATTAT 120  
 40 GCAATCATTT GTGCMCAAAA TAATGATCAG CAGGTCGAT TATTACTTTA TAATACTGAA 180  
 GATATAATTA ATTGGCATT TTTAGGTGAA ATAAATACAG AGTTGGATGA TTTTGGATAT 240  
 ATGTGGGAAT GCCCAGATTA CTTTAAATGA GATAATCAAG ATGTCATACT TATTTGTCCA 300  
 45 CAAGGTATTG AACCAAAAGG CGATCAGTTC AAAAATATTT ATCAAAGTGG TTATATACCT 360  
 GGAAGTTTGG ATATTGAAAA GTTAACATAT GAACATGAAA ATTTTGTGCA GCTTGATAAT 420  
 50 GGTTTTGATT TCTATGCACC TCAACATTT TTAGATGAAA AAGGCCGACG AGTACTAATT 480  
 GGATGGATGG GGTTACCGGA AATCGAATAT CCTACTGATA ATGAAGGATG GGCCCATTCG 540

GCGTTGGAAA AATTACGTCA CAATAAAGAG ACAGCATTCA GGcACGCAA ATAAATTTAC 660  
 TCGAAAATTA CATCCGTATG AAGGTAACA GTATGAATTA ATCATAGATA TTTTGGATAA 720  
 5 TGATGCTACC GAAGTGTACT TTGAATTACG TACAICTAAG ACTTCTTCAA CATTAAATTGC 780  
 TTATAACAAG CGTGGAAAATA AAATAACATT AGATCGCAGC GACAGTGGTT TATTGCCGAC 840  
 AAATGTTGAA gGTACGACGC GTAGTACGAT ATTAGACACG CCATTAA 887

10 (2) INFORMATION FOR SEQ ID NO: 476:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1183 base pairs  
 (B) TYPE: nucleic acid  
 15 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 476:

20 TTGGAAnCAA AAAACCATTC GTAACCGTG TtnAACCGGA TTCCGATGG ACCTTTTAAA 60  
 AChACCAAAAT AGAAAGCTTT GATAAAAGGT AATTATGGTA CTGATAACAA ACAAGTTCAA 120  
 25 AAACATCATG ATTTAGTACG TATGCTTTTG ATGGATCAAG ATGGTTTTTT AACTGAAAAAT 180  
 AATAAAGTTG ATCATTTCAT TGATGGAAAT GATTTATATG ATCAAGTTTT AAAAGATATT 240  
 AAAAAAGCAA AAGATATATC CATTTAGAGT ACwATACTTT CGCTTwAGAT GGTTWAGGTA 300  
 30 AAAGAATTTT ACATGCTTTA GAAGAAAAAT TGAACAAGG TCTAGAAGTA AAAATATTAT 360  
 ATGATGATGT TGGATCTAAA AATGTTAAGA TGGCAAATTT TGATCATTTT AAATCGTTAG 420  
 GTGGAGAAGT TGAAGCATTT TTTGCTTCAA AATTACCGTT ATTGAATTC AGAATGAATA 480  
 35 ATAGAAATCA TAGAAAAATC ATCGTAATCG ATGTCCACT AGGTTATGTC GGAGGATTTA 540  
 ACATTGGTGA TGAATATCTm GGATTAGGAA AATTAGGATA TTGGAGAGAT ACGCATTTAC 600  
 GTATACAAGG GGATGCGGTT GATGCACTGC AGTTGCGATT TATTTTAGAC TGGAAATCGC 660  
 40 AAGCGCACCG TCCACAATTT GAATATGATG TTAAGTATTT CCCTAAAAAG AACCGACCAT 720  
 TGGGCAATTC ACCAATTCOA ATAGCTGCAA GTGGCCCGCG TAGTGACTGG CATCAAAATTG 780  
 45 AATACGGTTA TACAAAAATG ATTATGAGTG CAAAGAAATC TGTATATTTA CAATCACCAT 840  
 ATTTCAATCC GGATAATTCA TATATAAATG CCATTAAAAT TGCTGCTAAA TCAGGTGTAG 900  
 ATGTACATTT AATGATTCCA TGTAAGCCAG ATCATCCATT AGTATATTGG GCGACATTTT 960  
 50 CAAATGCCTC TGACTTATTA TCAAGTGGTG TTAAAAATTA TACGTATGAA AATGGATTTA 1020  
 TACATTCTAA AATGTGCTTA ATTGATGATG AAATCGTATC AGTGGGCACA GCAAAATATGG 1080

55

EP 0 786 519 A2

CTAAAGATTT AAGGGTGGCT TATGAACATG ATATTACAAA ATC

1183

(2) INFORMATION FOR SEQ ID NO: 477:

- 5 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 2332 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 477:

|  |      |
|--|------|
| GGTTATATGT TTGGTATTAC TCATTATATT GATTGGAGGT TGTGTCATTA TGACAAAAAC      | 60   |
| 15 AAATGGTCGA AACGCTCAA TTAAGAAAAA TTCAACAAA ACATTAAAGTG TATATCTAAC    | 120  |
| CAAGAATCTC GATGATTTTT ACGATAAGGA AGGTTTTCGA GATCAAGAAT TTGATAAAAG      | 180  |
| 20 AGATAAAGGG ACTTGGAITA TTTATTCTGA AATGGTTATC GAACCAAAAG GGAAnAATAT   | 240  |
| GGAATCGAGA GGAATGGTGC TCTATATCAA TCGCAATACT AGAACmACGA AGGGTAATTT      | 300  |
| TATTGTCCAC GAAATAACTG AAGATAGTAA AGGATATICA CGTAGTAAAG AAAAAAATA       | 360  |
| 25 TCCTGTCAAG ATGGAAAAATA ATCGAATTAT TCCAACAAAG CCTATACCGG ATGACAAAGTT | 420  |
| AAAAAAGAG ATTGAAACT TTAAGTTCTT TGTACAAAT GGAATTTTA AAGATTTTAA          | 480  |
| AGATTATAAA AATGGTGATA TTTCATATAA TCCTAATGTG CCAAGTTATT CTGCAAAAGTA     | 540  |
| 30 TCAATTGAAT AATGATGACT ATAATGTTCA ACAGTTAAGA AAACGATATC ATATTCCAAC   | 600  |
| CAAAACAAGCG CCCGAATTAA AATTGAAAGG ATCCGGCAAT TTTAAAGGCT CATCCGTAGG     | 660  |
| ATCTAAGGAT CTAGAATTTA CGTTTGTAGA AAATCAAGAA GAGAATATCT ATTTTTCAGA      | 720  |
| 35 TTCGGTCGAA TTACACCTA GCGAGGATGA TAAATCATGA GTCAAACGGA ATATCAAATA    | 780  |
| AAATCTGGCA ATATAAAAGG TAACTCTGAA GAAACAAGTA CAGTATCTAA TATAAGTTAT      | 840  |
| GAAATAGAAA ACGCAATAA CAGTGGTTTA AAACAAAATA AAATTGATAA ACAAAATAAA       | 900  |
| 40 AAGTTACAAG AAAAAAATAA ATTCCTTAAA AATCTTTTCA ATCTTAAAG TTATACGGAC    | 960  |
| CCCCAACAG GCACGACTAC AAGCGCCTTT TTAAATAAAG ACACTGGCAA AGTTACTTTA       | 1020 |
| 45 GGTATGACAG GTACTAATGT ACACAAAGAC GCAATATTAA AACAAACATT TGGTGTTCCT   | 1080 |
| TCTTATCAAG GATATATAGA CGTGAGTGAA AcgCTAAAG ATATTGGGGC CGATGTCAAT       | 1140 |
| ATTGGCCTTC ATTCCGTAC AGATAAAGAT CCACATTATA AAAATACCCA AGACTTTATC       | 1200 |
| 50 AAAAATATCA AAAAAGACTA TGATATTGAT ATTATTACCG GACATTCGCT GGGCGGTAGA   | 1260 |
| GATGCGATGA TTTTAGGTAT GAGTAATGAT ATTAACATA TCGTTGTGTA TAATCCAGCT       | 1320 |

55

ATTGAAAAGT ACGATGGTCA CATGTGAAGA TTTGTGCTG ATGAAGACGA ATTAGATGCA 1440  
 GGTGTCCGCA ATCATTTATA TGAAGCTGCT GGAGAAAAAA TAGTACTTAA AAATGGAGAA 1500  
 5 GGCCATGCAA TGAGTGGTAT TTTAATGAGC AGAACACAGG CTATAATCTT AGCTGAATTA 1560  
 AACAAAGTTA AAGGCTACCA AGACGAAAAA AATAAAGCAT TAAATCCGT TCGTAAACAA 1620  
 ACGAGGCATA GATTACATAA AGTAGAGACG TTAAGAGCGA ATTGATTTCA AACACGGGT 1680  
 10 GGATCACTCT CTTCCTCCW ACAACAATTA TTAGAAGCTT TAAACAGACT AACCATTTGCC 1740  
 GAAGGCTTAA ATCAATTAGT GAATGAAGAA AGCCACACTT TGAAAAAATG TATCACGCGA 1800  
 TGGCACATAA ATTTGGAGAC AACTGGAAAA AAGCGCAAGA AGTTGGAATG GAAATTGGTG 1860  
 15 AAAAAATTA CTCTGAAGAG GTTATAGATG FATTAAAGAA AGGTGGCGCG TATGAAAGTa 1920  
 AACTTGAAC AGATCCCAA AGAAAAATTG ATGATAAGAT AAAGAAATTA AATGATGTTT 1980  
 ATAAAAATTG TAATGGCTAT ATCGCAAAAA TTAACAGAG TATCGAAGCA ATTGTTTCTA 2040  
 20 ATGACCAAT GTTAGCGAGC CAGATTGATG GGATGATGTA ATGTTTACTA CGTATAAnAA 2100  
 TATTATAGAA CTGAAAAATG CCTATGATGA AGAAAGAAAA CAATTGAATG ATGCATTTCAA 2160  
 25 TCAAAATGAT GAATTAAGAC ATCAAAACAG CAAGAmATGT GAACAAATGT ATGATCATTT 2220  
 CTTATATCTC AAACATAAAA TGAATTmym TGAAGACGCT ATGATCAGGA TGACACGTAT 2280  
 TATAGAATCT TTCGATAGAG AAACGAATCA ACGTATCCGA CATCACGAAA TG 2332  
 30 (2) INFORMATION FOR SEQ ID NO: 478:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 865 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 35 (D) TOPOLOGY: linear  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 478:  
 40 TTACTACCC AGTATCTCTT TTTATAAAAT ATATAGCCAC CACATATGGT GGAAAGTCTT 60  
 TTTAATTAGA ATTTGTTTTT TTCAGTTAAG AAAGCTTCTA ACTCTGAGAT TGGCATACGA 120  
 45 ACTTGTTCCTA TTGAATCTCT GTCAGTACT GTAACCTGAT TATCTTCTAA TGAATCAAAG 180  
 TCGAATGTtA CACAATAAGG TGTtCCGATT TCATCTTGAC GACGGTATCT TTTACCGATA 240  
 GATGTGATT CAICGAAATC GATTGAGAAT TTAGAACTTA ATTGCTCAA AATCTTAATC 300  
 50 GtTCGCCAGA TAATTTCTTA CTAAAGGTA AAATCGCTGC TTTATATGTT GCTAATGCAG 360  
 GATGGAAGTG TAAAACTGTA CGTGATCTTT TACTACCTTC AACGCCTTCT TCATCATATG 420

55

GAATATATTT TCGTTCGTT TCTGGATCAT GGTATCTGAA ATCTTCACCA GAGTGTTCAG 540  
 CATGTTTACG TAAGTCGAAG TCTGTACGAC TTGCGATACC CCATAACTCA CCCCACCCAA 600  
 5 ATGGGAATTT ATATTTCAATA TCAGTTGTGT CATTGAGTA ATGAGATAAT TCATCTTCAT 660  
 CATGATCACG TAAAGCATA TTTTCACTGC TCATATTTAA GCTTGTTAAC CAGTCACTTG 720  
 CAAAAGTTTT CCAATAATTT TGCCATTGCA TTTCTTCTCC AGGTTTACAG AAGAATTCAA 780  
 10 GTTCCATTGG TTCAAATTTCT CTGTGTTCTGA AAATGAAGTT ACCTGGAGTG aTTTCaTTAC 840  
 GGAATGaTTT ACCAATTGG ACCGG 865

(2) INFORMATION FOR SEQ ID NO: 479:

15 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1444 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 20 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 479:

25 TAATGAGTAA ATAAGTACCA GATAAAATCA TGAATATCAT CCAACATGCG GTTAACTCTA 60  
 CTAATAATT AATAAGTGA TTTTCAGTAA ACAAGAAGCT ATGTATACCT CGCATCACAT 120  
 TAGAATACGT ATGTTTCGCA TTTTGATCTG CAACAAATG ATTGTTATGA TCTAGAAAAGA 180  
 30 CGTAACGTTG ATTTCTGCGC ATATCACTCA GTGTAATTGG TTGTTATAT GTTTCATCAA 240  
 GTATGCTAAC TTTACTTACA AAGAATCCCT CATATTGTTG TTCAACTTGA TGTTACTGCAT 300  
 CATTAAATGT TTGATGCGTT TTTACATCAC TGTACCAAAA AAACATCATC TTATAAATAT 360  
 35 TATTTTCAAC TTCTGGAAAG AACAAAGTAAC CAATGCCCGA AATGGTTAAA GTGATTAACA 420  
 GTGAGCAAT AAATATTGCT GCATAGAAAT GTAATCTTTG TAATGGATTA AATGATTTTT 480  
 TCATATTTCC CTCCCAATTG GCTATTATAC GGTGTCAATT CTGTGATGTG TGTGAACAAA 540  
 40 CTGTGACAA ATTTATTTTC TAGAAAAAAT TAACGATGAT TTGTGATTTT TAGAAAAATG 600  
 AACTTTTAAG TTGGAATGTT TGAAGAAAAA TGATTATTCG TATGTTTTAT CAAGCAGCTA 660  
 TGATAAAATT TAAACATAAT ACAATGCGAG CCATTTAAGC ATCTATGTTT AAATGGACAT 720  
 45 CGATATTGTA TGAATTCGTT GTAACAAGCA AGCATTTCTA TGTGAACGAA CCAAGGGGGA 780  
 AAGTAACATG ATTAATAAAG AACAAATAGA TCTTTTATAT AAATTAAGAA AAGAAGTTGA 840  
 50 AAGTCGCGA AATGAAGCAC TTTTACATAC AATTAACCAA GTAATTAAGA AAGTATATTT 900  
 GCAGCAATAT ACATGTTCTG TCGTTGACAA TTTTCTGCA GGTAAATCGA CACTGATAAA 960

55

## EP 0 786 519 A2

TATTGTGICA GTTTCAGACA ATCAGCATAT TATTGCTAAT TTGCCGAATC AAACGTATGC 1080  
 CAAATTATCT AATTATGATG AAGTAAGGGA AATGAATCGC CAAAATGTGC ACGTTGAATC 1140  
 5 TGTGAAAAAT AATTTTCAAT CAGCTAAATT TGAAAAATGGG TTTACGTTGC AAGATACACC 1200  
 AGGTGTTGAT TCAAATGTTG CATCACATCA GTCAATAACA GAACAATATA TGTATACAAG 1260  
 TAATATGATA TTTTATACGG TTGACTATAA CCACGTTCAA TCTGAACCTA ACTTTAAGTT 1320  
 10 TATGAAGCAT ATAAATGATG TTGGAATACC TGTGTGTTT ATCATTAAATC AAATTGACAG 1380  
 CATCCAAGAC GATGGAATTG TCATTCTCTA CGTnTAAAT CTCGAGTTGG AAAAACTAAT 1440  
 TGGC 1444

15 (2) INFORMATION FOR SEQ ID NO: 480:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 6309 base pairs  
 (B) TYPE: nucleic acid  
 20 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 480:

GCAGCAGCTT ATCGTGAGTT ATCATTATTA TTACGTAGAC CTCAGGTCG TGAACATAC 60  
 CCAGGTGACG TATTCTACTT ACATAGTAGA TTATTAGAAA GAGCAGCAAA ATTAACGAT 120  
 30 GACTTAGGTG GCGGTTCAAT TACTGCATTA CCAATTATCG AAACACAAGC TGGTGATATT 180  
 TCAGCTTATG TACCAACAAA CGTTATTTC AATTACAGAT GACAAATCTT CTTACAATCT 240  
 GATTATATCT TCTCAGGTGT AAGACCAGCG ATTAATGCCG GACAAATCTGT ATCTCGTGT 300  
 35 GGTGATCTG CACAAATTAA AGCAATGAAG AAAGTTGCTG GTACGTTACG TCTTGACTTA 360  
 GCGTCATACA GAGAACTTGA ATCATTTGCA CAATTGCGTT CAGACCTGA TGAATTACT 420  
 GCAAGTAAAT TAGAACGTGG TAAACGTACT GTTGAAGTCT TAAACAAGA TCAAAACAAA 480  
 40 CCATTACCAG TCGAACACCA AGTGTTGATT ATTTATGCAT TAACAAAAGG ATATTAGAT 540  
 GATATTCAGT TTGTAGATAT CACACGTTTT GAAGACGAGT TAAACCACTG GGCAGAAATCA 600  
 AATGCTACTG AACTGTTAAA TGAAATCAGA GAACTGGTG GCTTACCAGA TGCTGAGAAG 660  
 45 TTTGACACAG CAATTAAACGA ATTCAAAAAA AGCTTTAGCA AATCTGAATA ATAAACAAGT 720  
 TTAGATAAAG GTGGTGAGAT AGTGCTTCT CTTAAAGAAA TAGATACTGG AATAAAATCA 780  
 50 ACCAAAAAAA TGAAGCAGAT TACGAAAGCG ATGAACATGG TATCAAGTTC AAAACTCGT 840  
 AGAGCTGAAA AAAATACAAA ACAATTACCA CCATATATGG ATAAATGCA AGATGCAATT 900

55

|    |             |            |             |             |             |             |      |
|----|-------------|------------|-------------|-------------|-------------|-------------|------|
|    | ACTAGAAGTG  | GATATTTAGT | TATCACGAGT  | GATAAAGGTT  | TAGCAGGTGC  | ATATAGTGCA  | 1020 |
|    | AACGTGCTTA  | AAAAATTGAT | TACTGATATT  | GAAGCGAAAC  | ATCAAGATAG  | TAGCGAATAC  | 1080 |
| 5  | AGTATTGTAG  | TTTTAGGGCA | ACAAGGTGTT  | GATTTCCTAA  | AAAATAGAGG  | TTATGACATT  | 1140 |
|    | GAGTATTCTC  | AAGTAGACGT | ACCTGATCAA  | CCTTCTTTCA  | AATCTGTTCA  | AGCACTAGCT  | 1200 |
|    | AACCATGCTA  | TAGACTTATA | CAGTGAAGAA  | GAAATTGATG  | AATTAATAT   | ATACTATAGT  | 1260 |
| 10 | CATTATGTCA  | CGTTCCTTGA | AAACAAGCCT  | ACATCTAGAC  | AAGTATTACC  | ATTATCTCAA  | 1320 |
|    | GAGGATTCTA  | TGAAGGGGCA | TGGTCATTTG  | TCTTCTTAGT  | AATTTGAGCC  | AGATAAAGAA  | 1380 |
|    | TCTATTCTTAA | GTGTAATCTT | GCCTCAATAT  | GTTGAGAGTT  | TGATTTACGG  | AACAATTATTA | 1440 |
| 15 | GACGCAAAAG  | CAAGTAGGCA | TGCAACACGT  | ATGACCTGCGA | TGAAAATATG  | CACGTGTAAT  | 1500 |
|    | GCAACTGAAC  | TTATTGATGA | CTTATCATT   | GAATATAACA  | GAGCGAGACA  | AGCAGAAATT  | 1560 |
|    | ACGCAACAAA  | TTACTGAAAT | TGTTGGTGGT  | TCCGACGCG   | TTGAATAATA  | TTTAAAGGAG  | 1620 |
| 20 | GAAAATAGCA  | TGGGAATTGG | CCGTGTAAC   | CAAGTTATAG  | GTCCTGTAAT  | TGATGTTCCA  | 1680 |
|    | TTTGAACATA  | ACGAAGTTCC | TAAAAATTAAT | AACGCCCTTG  | TTATTGATGT  | GCCTAAGAA   | 1740 |
| 25 | GAGGTACAA   | TACAACATAC | ATTAGAAGTT  | CGCGTGCAAT  | TAGGTGACGA  | CGTGTTCGT   | 1800 |
|    | ACAATTGCGA  | TGGAATCAAC | TGATGGTGT   | CAAGAGGGCA  | TGGATGTAAA  | AGATACAGGC  | 1860 |
|    | AAAGAAATTA  | GTGTACTCGT | TGGTGACGAA  | ACATTAGGTC  | GTGTATTTAA  | TGTACTAGGT  | 1920 |
| 30 | GAAACAATTG  | ACCTTAAAGA | AGAAATTAGT  | GATTCCTGTC  | GCCGCGATCC  | TATCCATCGT  | 1980 |
|    | CAAGCACCAG  | CATTCCGATG | ACTTTCACAA  | GAAAGTCAAA  | TTTTAGAAAC  | AGGTATTAAA  | 2040 |
|    | GTAGTAGATT  | TACTAGCACC | TTATATTAAA  | GGTGTAAAA   | TGGGATTGTT  | CGGTGGTGCC  | 2100 |
| 35 | GGTGTAGGTA  | AAACAGTATT | AATCCAAGAA  | TTAATTACAA  | ACATCGCTCA  | AGAGCAGCGT  | 2160 |
|    | GGTATTCTGT  | TATTCCCGGG | TGTAGGTGAA  | CGTACTCGTG  | AAGGTAAACGA | TTTATACTTC  | 2220 |
|    | GAAATGAGTG  | ACAGTGGTGT | AATTAAGAAA  | ACAGCCATGG  | TATTGGGGA   | AATGAATGAG  | 2280 |
| 40 | CCACCTGGTG  | CACGTATGCG | TGTTGCATTA  | TCTGGTTTAA  | CAATGGGCTGA | ATATTTCCTG  | 2340 |
|    | GACGAACAG   | GTCAAGACGT | ATTATTATT   | ATCGATAACA  | TTTTAGAGT   | TACACAAGCT  | 2400 |
| 45 | GGTTCGAGG   | TATCTGCATT | ATTAGGTCGT  | ATGCCCTCTG  | CAGTAGGTTA  | CCAACCAACA  | 2460 |
|    | CTGTCTACTG  | AAATGGGACA | ATTACAAGAA  | CGTATTACGT  | CTACAACAA   | AGGATCAGTT  | 2520 |
|    | ACTTCTATT   | AAGCGGTATT | CGTACTGTCC  | GATGACTATA  | CTGACCAGC   | GCCTCGAGCA  | 2580 |
| 50 | CGGTTTGCC   | ATTTAGATGC | AACTACAAAC  | TTAGAACGTA  | AATTAACCTA  | AATGGGTATT  | 2640 |
|    | TATCCAGCGC  | TGGATCCATT | AGCGTCTACA  | TCAAGAGCAT  | TGGAACCATC  | AATTGTAGGT  | 2700 |



|    |  |      |
|----|--|------|
|    | CAAGATATCA TTGCTATCIT AGGTATGGAC GAATTATCTG ATGAAGATAA ACAAACAGTT  | 2820 |
|    | GAACGCGCAC GTAGAATTCA ATTCCTCTTA TCTCAAACT TCCACGTAGC GGAACAATTT   | 2880 |
| 5  | ACTGTCAAA AAGGTTCTTA TGTACCTGTT AAGACAACAG TTGCAAACTT TAAAGATATC   | 2940 |
|    | TTAGATGGTA AATATGACCA TATCCAGAA GATGCATCC GTTTAGTTGG TAGCATGGAT    | 3000 |
|    | GATGTTATTG CAAAAGCTAA AGATATGGGT GTTGAAGTAT AACAAATTAG AGGAATGGAT  | 3060 |
| 10 | AATGAATACA TTAAACCTAG ATATGTGCAC TCCTAATGCT TCTGTTTACA ATCGTGATAA  | 3120 |
|    | TGTTGAACTC GTTGTATATG AAACAACAGC TGGTGAGATA GGTGTCATGA GTGGACATAT  | 3180 |
|    | TCCAACGTGA GCTGCTTTAA AAACAGGCTT TGTAAAAGTG AAATTTACAG ATGGAACTGA  | 3240 |
| 15 | ATATATTGCT GTAAGCGATG GCTTTGTTGA AGTTAGAAAA GATAAAGTTT CAATCATTGT  | 3300 |
|    | TCGAGCTGCA GAACTGCAA GAGAAATTGA TGTGAAAGA GCTAAATTAG CCAAGCAAG     | 3360 |
|    | AGCAGAGTCT CACTTGGAAA ATGATGACGA CAATACTGAT ATTCATAGAG CCGAAAGAGC  | 3420 |
| 20 | TTTAGAGAGA GCAAAATACC GTTTCGTGT GCGTGAATTA AAATAGTAAA TAAAGGGTCG   | 3480 |
|    | AAGATGTGAT TTCATATCTT CGACCCCTTT TTGAATTATA TTGATTTAAA GATACAAAAC  | 3540 |
| 25 | ATGAGAGGGG GGAAGGAATT GATAAAGAAC CATTAAAGAT TTATGATGA GTGGTCTCTT   | 3600 |
|    | ATCATTAAAC ACAGCTAATG TGTATTTAAA AATAGGaaYa CATGAGTAAA ACTCATGTAT  | 3660 |
|    | AAGAAATACT AATTTCTAAA GAAAAAGTAT TTCTTTATGT TGGGGCCCCG TCAACTACTG  | 3720 |
| 30 | CCAAATACAA CACTATAGAG TCTAGACATT GATTTATGTC CGACTCCCAA GAATAGTTTT  | 3780 |
|    | ACTTTTTTAC AATCACTAAT AGATTGCTAA AATCAAAATT TCCTTCACCA CTATCTACAG  | 3840 |
|    | TCGACATTC ATTTTTGAA ATTATCTACA TTTTTCATA CCAAGATATT TTATAGTTAT     | 3900 |
| 35 | GATATTTATG TAAAAAGAAT TATATAGTAA GTTAGCTTAA ACTTTACTAA AAACGGGTAT  | 3960 |
|    | TAAACTTTGT ATCATTATTT AAATTTTTCA TGTACAATGT AATACAGTAA TCTTATGAGG  | 4020 |
|    | TGATAAAATG GATTATATCG GACAATATGC AGTTATCCAT TTAGTGTATC ATGTTGTATG  | 4080 |
| 40 | TATTTGTATT GCCTATTGGG CTTTACAATC AATTAGATTA GATCAATTTT TTAaaaaaAG  | 4140 |
|    | ATACGCCACT CAATTACAAG TGTGTATGAT ATTTGTTGCT ATTTTATTAG GCACGTGCACT | 4200 |
| 45 | AAGCAATTTT ATTGTAGATT TGTACAATA CTCGACGCG AGTAAAAATT TAATAAAATA    | 4260 |
|    | AGTCTAACTC TATGATTGTG AATCAAACT AGATATAATT AAATAATGAC TTAATAATAT   | 4320 |
|    | TTTAAATAG GGAAGGTAA AGTAATAGGA GTTCTAAGTG GAGGATTTAC GATGGATAAA    | 4380 |
| 50 | ATAGTAATCA AAGGTGAAA TAAATTACG GGTGAAGTTA AAGTAGAAGG TGCTAAAAAT    | 4440 |
|    | CGAGATTAC CAATATTGAC AGCATCTTGA TTAGCTTCTG ATAAACGAG CAAATTAGTT    | 4500 |

|    |             |            |             |             |            |             |      |
|----|-------------|------------|-------------|-------------|------------|-------------|------|
|    | GACGTTACAT  | ACAAAAAGGA | CGAAATGCT   | GTGTCGTTG   | ATGCAACAAA | GACTCTAAAT  | 4620 |
|    | GAAGAGGGCAC | CATATGAATA | TGTTAGTAAA  | ATGCGTGCAA  | GTAATTTAGT | TATGGGmCCT  | 4680 |
| 5  | CTTTTAGCAA  | GACTAGGACA | TGCTATTGTT  | GCATTGCGTG  | GTGGTTGTGC | AATTGGGAAGT | 4740 |
|    | AGACCGATTG  | AGCAACACAT | TAAAGTITTT  | GAAGCTTTAG  | CGCGAGAAAT | TCATCTTGAA  | 4800 |
|    | AATGGTAATA  | TTTATGCTAA | TGCTAAAGAT  | GGATTAAAAG  | GTACATCAAT | TCATTTAGAT  | 4860 |
| 10 | TTTCCAAGTG  | TAGGAGCAAC | ACAAAATATT  | ATTATGGCAG  | CATCATAGC  | TAAGGGTAAG  | 4920 |
|    | ACTTTAAATTG | AAAAATGCAG | TAAAGAACCT  | GAAATTTGCG  | ATTTAGSCAA | CTCATATTAA  | 4980 |
|    | GAAATGGGTG  | GTAGAATTAC | TGGTGCTGGT  | ACAGACACAA  | TTACAATCAA | TGGTGTAGAA  | 5040 |
| 15 | TCATTACATG  | GTGTAGAACA | TGCTATCATT  | CCAGATAGAA  | TTGAAGCAGG | CACATTACTA  | 5100 |
|    | ATGCGTGGTG  | CTATAACGGG | TGGTGATATT  | TTTGTACGTG  | GTGCAATCAA | AGAACATATG  | 5160 |
| 20 | GCGAGTTTAG  | TCTATAAACT | AGAAGAAATG  | GGCGTTGAAT  | TSGACTATCA | AGAAGATGGT  | 5220 |
|    | ATTGCTGTAC  | GTGCTGAAGG | GGAATTACAA  | CTGTAGACG   | TCAAAACTCT | ACCACATCCT  | 5280 |
|    | CGATTCGCCG  | CTGATATGCA | ATCACAAAAT  | ATGGCATTTG  | TATTAACGGC | AAATGGTCAT  | 5340 |
| 25 | AAAGTCGTAA  | CGGAAACTGT | TTTTGAAAAC  | CGTTTATATG  | ATGTTGCAGA | GTTCAAACGT  | 5400 |
|    | ATGAATGCTA  | ATATCAATGT | AGAAGGTCGT  | AGTGCCTAAC  | TTGAAGGTAA | AAGTCAATTG  | 5460 |
|    | CAAGGTGCAC  | AAGTTAAAGC | GACTGATTTA  | AGAGCAGCAG  | CCGCGTTAAT | TTTAGCTGGA  | 5520 |
| 30 | TTAGTTGCTG  | ATGTTAAAC  | AAGCGTTACT  | GAATTAACGC  | ACCTAGATAG | AGCGTATGTT  | 5580 |
|    | GACTTACACG  | GTAATAAGAA | GCAATTTAGT  | GACATAGTAA  | ACCGTATTAA | CGAATTAAATC | 5640 |
|    | AGTAAATTAA  | TATAATTGAG | GATTTCAACT  | ATGGAACCAA  | TTTTTGATTA | TAACCAAAAT  | 5700 |
| 35 | AAACAAATTA  | TACCTCACAG | ACAGCCATTT  | TTATTAATTG  | ATAAAGTAGT | TGAATATGAA  | 5760 |
|    | GAAGGTCAAC  | GTGTTGTGGC | TATTAACCAA  | GTATCAGGAA  | ACGAACCAT  | CTTTCAAGGG  | 5820 |
| 40 | CAITTTCCCTG | AGTATGcGGT | AATGCCAGGC  | GTAATTAATTA | CTGAAGCGTT | AcTCAAAACAG | 5880 |
|    | GTGCGGTAGC  | TATTTTAAAT | AGTGAAGAAA  | ATAAAGGTAA  | AATCGCTTTA | TTTGTCTGTA  | 5940 |
|    | TGATAAAATG  | TCGTTTTTAA | CGTCAAGTAG  | TACCTGGTGA  | TACTTTAACG | TTGGAAGTAG  | 6000 |
| 45 | AAATCACTAA  | AATTAAGGAA | CCAATAGGTA  | AAGGTAATGC  | TAAAGTACT  | GTGATGGTC   | 6060 |
|    | AACCTGCTTG  | TAGTTGTGAG | CTTACATTGG  | CAATTCAAGA  | TGTAAAAATA | AACAAAAAAA  | 6120 |
|    | ACATTCAAAAG | ATTTAATGTG | TTGGCAGTAAT | CTTTGAATGT  | TTTTTATTTT | ACTCTTCTAA  | 6180 |
| 50 | TTTTTCTATC  | TTTAACITTT | CTTTAGATCG  | CATCATTCGA  | TAAATGATTT | TTTTTAATTC  | 6240 |
|    | TTCAACGAGAT | AATCCATCAT | CAATAAGTTG  | GTTCATAATA  | ACTTTACAGA | TACTGTTGGA  | 6300 |

EP 0 786 519 A2

(2) INFORMATION FOR SEQ ID NO: 481:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1717 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 481:

|             |            |             |            |             |            |      |
|-------------|------------|-------------|------------|-------------|------------|------|
| AGTTGCTACA  | CCAGACATGA | TGGGTGAAGT  | TGGTAAATTA | GGTCGTGTAT  | TAGGACCAAA | 60   |
| AGGTTTAAATG | CCAAACCCTA | AAACTGGAAC  | TGTAACAATG | GATGTTAAAA  | AAGCTGTGTA | 120  |
| AGAAATCAAA  | GCTGGTAAAG | TAGAATATCG  | TGCTGAAAAA | GCTGGTATCG  | TACATGCATC | 180  |
| AATTGGTAAA  | GTTTCATTTA | CTGATGAACA  | ATTAATtGAA | AACyTcAATA  | CyTTACAAGA | 240  |
| TGTATTAGCT  | AAAGCTAAAC | CATCATCTGC  | TAAAGGTACA | TACTTCAAAT  | CTGTGTCTGT | 300  |
| AACACAACA   | ATGGGTCCTG | GAGTTAAAAAT | TGATACTGCA | AGTTTCAAAT  | AATAAATGAT | 360  |
| ATAACAATT   | ACAGGCTGAA | AGAAATATCT  | TTCACTCTGT | AAAAATATAT  | TGCAATATAG | 420  |
| TAATTTCCAA  | GTTATATTAC | TTATTGTGAT  | TATTTTACCT | ANGACAGTAG  | GAGTTATTTA | 480  |
| TAACCTAAAA  | TTTATCCTGC | CGAGGCTAAA  | ATTGACTTGA | ACGTGATGAT  | CTATGATCTT | 540  |
| TCAAGCACTT  | TTTGCCGTGG | GTAGAAAGTG  | CTTTTTTTAT | TAATTTTAAA  | AAAAGCACCA | 600  |
| AAAATTTAAA  | TGGAGGTGTC | TGAATGTCGT  | CTATCATTGA | AGCTAAAAAA  | CAACTAGTTG | 660  |
| ATGAAATTGC  | TGAGGTACTA | TCAAATTCAG  | TTTCAACAGT | AATCGTTGAC  | TACCGTGGAT | 720  |
| TAACAGTAGC  | TGAAGTTACT | GACTTACGTT  | CACAATTACG | TGAAGCTGGT  | GTTGAGTATA | 780  |
| AAGTATACAA  | AAACACTATG | GTACGTCGTG  | CAGCTGAAAA | AGCTGGTATC  | GAAGGCTTAG | 840  |
| ATGAATTCCT  | AACAGGTCCT | ACTGCTATTG  | CAACTTCAAG | TGAAGATGCT  | GTAGCTGCAG | 900  |
| CGAAAGTAAT  | TTCTGGATTI | GCTAAGATC   | ATGAAGCATT | AGAAATTAAA  | TCAGGCGTTA | 960  |
| TGGAAGGCAA  | TGTTATTACA | GCAGAAGAAG  | TTAAAACTGT | TGGTTCAITTA | CCTTCACACG | 1020 |
| ATGGTCTTGT  | ATCTATGCTI | TTATCAGTAT  | TACAAGTCC  | TGTAGGCAAC  | TTGCTTATG  | 1080 |
| CGGTAAAGC   | TATTGGAGAA | CAAAAAAGAAG | AAAACGCTGA | ATAATTTTITA | CGGTAAAAAA | 1140 |
| ATTAAAAATA  | ATGGAGGAAT | TATAAAATGG  | CTAATCATGA | ACAAATCATT  | GAAGCGATTA | 1200 |
| AAGAAATGTC  | AGTATTAGAA | TTAAACGACT  | TAGTAAAGC  | AATTGAAGAA  | GAATTTGGTG | 1260 |
| TACTGcAGCT  | GCTCCAGTAG | CAGTAGCAGG  | TGCAGCTGGT | GGCGCTGACG  | CTGCAGCAGA | 1320 |
| AAAAACTGAA  | TTTGACGTTG | AGTTAACTTC  | AGCTGGTTCA | TCTAAAAATCA | AAGTTGTGTA | 1380 |

|   |  |      |
|---|--|------|
|   | TCCTAAAGTA ATCAAGAAG CTTTACCTAA AGAAGAGCT GAAAACCTTA AAGACAATT   | 1500 |
|   | AGAAGAAGTT GGAGCTACTG TAGAATTAAA ATAACTCAAG TATCTTAAAC TTAATATCA | 1560 |
| 5 | AAGTTTTATA GCAAGTATTG CTATAATATA ATGTTCTTT GAGAAGTTAA AACCCCGTTA | 1620 |
|   | TTTTGATAAC GGGgTtTTT TcATTTAAAG ACTGAGTGAC ATGTTTAAAT TATAATGAC  | 1680 |
|   | AGTTACAAGG TGAAGATGAG GTGGGAACTA TAGTCA                          | 1717 |

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1279 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

|    |            |            |            |            |             |             |      |
|----|------------|------------|------------|------------|-------------|-------------|------|
|    | GTAATCTGTT | TATCGGAAT  | GTAACGATTA | AAGAAAAGGG | CTCATCTCAA  | ACATATATTT  | 60   |
|    | TGTTAGGCTA | TCCACAACAA | GCACAGAAGA | ATAGTCATAG | CAAAATATAG  | GGAGTCTTTA  | 120  |
| 25 | TATATAAAGA | CTTGAATCA  | ATCGAAGATA | CAAAATATGC | TATTACGATT  | ATCCACATAA  | 180  |
|    | TTACGGCTGT | TATTTTCTTA | ACAATTACAA | CAGTCTTTGT | GTTTTTCTTA  | TCGTCAGAGAA | 240  |
|    | TTACAAAACC | TTTAAGACGT | ACGAAGAGAC | AAGTCACACG | TGTATCTGAA  | GGGGATTACT  | 300  |
| 30 | CTTATAAAAC | TTCTGTCCA  | ACGAAAGATG | AAATTGGTCA | ATTATCGCAG  | GCATTTAATC  | 360  |
|    | AGATGAGTAC | AGAAATCGAA | GAGCATGTGC | ACGCATTATC | CACATCTAAA  | AAATTATAGAG | 420  |
|    | ACAGCTTAAT | TAACTCTATG | GTAGAAGGTG | TCCTAGGTAT | TAATGAGAGT  | CGACAATAAT  | 480  |
| 35 | TCTTATCTAA | TAAGATGGCG | AATGATATTA | TGGACAATAT | TGATGAAGAT  | GCTAAAGCTT  | 540  |
|    | TCTTATTAA  | ACAAATAGAA | GATACCTTTA | AATCAAAACA | AACTGAAATG  | CGCGATTTAG  | 600  |
|    | AAATGAATGC | ACGATTCTTT | GTTGTGACCA | CAAGCTATAT | CGACAAGATT  | GAACAGGGAG  | 660  |
| 40 | GTAAAGTGG  | TGTTGTTGTG | ACAGTTCGTG | ATATGACTAA | TAGACACATA  | CTAGATCAAA  | 720  |
|    | TGAAGAAAGA | TTTCATTGCT | AATGTATCAC | ATGAATTACG | TACACCGATA  | TCATTACTTC  | 780  |
| 45 | AAGGTTATAC | TGAATCAATT | GTAGATGGTA | TGTTACAGAA | ACCGGATGAA  | ATAAAAGAAT  | 840  |
|    | CGCTTGCCAT | TGTCCTTGAT | GAATCGAAAC | GTTTAAATCG | TTTAGTTAAT  | GAATTGTTAA  | 900  |
|    | A1GTGCGACG | CATGGATGCT | GAAAGGTTAT | CCGTAAATAC | AAGAAGTTACG | CCTATTGCAC  | 960  |
|    | CGTACTAGAA | TAGATGAAA  | ATTAAAGTAT | GCCACAACAG | TGATGATTTA  | GGTCTTAAATA | 1020 |
| 50 | TGACTTTTAA | TATGTTGAAG | AAGCGTGTGT | GGAGTTATGA | TAWGGATCGC  | ATGGACCAAG  | 1080 |

## EP 0 786 519 A2

TTACTTGTGA TGAAATGAA AGCGAAGATA TTTTATACAT TAAAGATACA GGACAGGCA 1200  
 TTGCACCAGA ACAATTACaA CAAGTATTtg ATCGTTTTTA TAAAGTTGAT GCAGCGaAnA 1260  
 5 ACCCGnGGT AACCAgTA 1279

(2) INFORMATION FOR SEQ ID NO: 483:

(i) SEQUENCE CHARACTERISTICS:  
 10 (A) LENGTH: 1144 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 483:

GAGCTGTTGT TACTTTGATG CCTGCAGCTT TATTACGGCT GACTTGGTAA TGATAAGTTT 60  
 20 CAGCATATTG CTCAATATAT GCTATATCAT ATTGAATGGT ACGAGGTGAT ACACCAAGTT  
 GATTAGCAAT GGTATTGATT GGAATAAACG TTGCTCATG AATTAAAAGA TACAAAAATT 180  
 CGATTTGCTC ATAACCTAAC AACGTAATAT CCTCCTATTT GTAATTGTAA GCGATTCTTT 240  
 25 AAAAAAGTAG ATATGCAATC TCTTTCATAT TTTAATCGA AAAATTGCAT ATCAAAATGT  
 TTATGGCGCA AGATTTTATA GGAACCTTTA AAATAAAITA xATATTCATG TTGACAAATT 360  
 AAAAAATGCG CAGTATAATT AGTTAGACAT CTAACGAAAT GGTGGTGCAA TAAATGGAAT 420  
 30 TCACTTATTC GTATTTATTT AGAATGATTA GTCATGAGAT GAAACAAAAG GCTGATCAAA  
 AGTTAGACGA ATTTGATATT ACAAAAGAGC AAGGTCATAC GTTAGGTTAT CTTTATGCAC 540  
 ATCAACAAGA TGGACTGACA CAAAATGATA TTGCTAAAGC ATTACAACGA ACAGGTCCAA 600  
 35 CTGTCAGTAA TTTATTAAGG AACCTTGAAC GTAAAAAGCT GAICTATCGC TATGTCGATG  
 CACAAGATAC GAGAAGAAAG AATATAGGGC TGACTACCTC TGGGATTAAA CTCGTAGAAG 720  
 40 CATTCACTTC GATATTTGAT GAAATGGAAC AAACACTCGT ATCGCAGTTA TCTGAAGAAG  
 AAAATGAACA AATGAAAGCA AACTTAACTA AAATGTTATC TAGTTTACAA TAAATGATAA 840  
 GTGTGACTGG TAGAAATCAG TCACTTTGTC TTTAATATTA TAGTTAGATA TCTAATTGTT 900  
 45 AGTAAGCTAA TTATTGGAAA AGACAAGGAG TATTGAACAA TGAAGACGA ACAATTATAT  
 TATTTTGAGA AATCGCCAGT ATTTAAAGCG ATGATGCATT TCTCATTGCC AATGATGATA 1020  
 GGGACTTTAT TAAGCGTTAT TTATGGCATA TTAATATTTT ACTTTATAGG ATTtTyAGam 1080  
 50 GAYAGCCACA TGATTCTGCG tAatCTCTCT AACACTGCCA GTATTTGCTA TCTTAATGGG  
 GTTA 1144

55

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1158 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

|    |            |             |             |             |             |             |      |
|----|------------|-------------|-------------|-------------|-------------|-------------|------|
| 15 | ACAACCAATT | TTACTAAACA  | TGGrTtTCAT  | gCATTTTCAA  | TATTTAGATG  | TCGACCGCAT  | 60   |
|    | AATCGAAGAA | TCGCGGACAA  | TAGTACTTAT  | CGATGAGTTA  | GCACATACGA  | ATATTCTTAG  | 120  |
|    | AGATCGTCAT | GAGAAACGAT  | ATATGGATTA  | TGAAGAAATT  | TAAATCATG   | GTATCGATGT  | 180  |
|    | TCATACCATT | TTGaACATTC  | aaCATATTGA  | AAGTTTAAAG  | AGTCAAAATT  | AACTCGATGAC | 240  |
|    | CGGTGTACAT | GTTAAAGAAG  | GTGTACCCGA  | CTATTTTATA  | ATGAGCGCCG  | ATGTATTAGA  | 300  |
| 20 | AGTCGTAGAT | ATCTCACCTG  | AAACAATTAAT | TAAACGCTTA  | AAAGCTGGCA  | AGGTATATAa  | 360  |
|    | AAAGGATAGG | CTAGATAGTG  | CATTTAGTAA  | TTTCTTTTACG | TATGCCCAAC  | TAAAGCGAAG  | 420  |
|    | CGTACATTGA | CGTTAAGAAC  | ATGTCGCGAC  | TGTGTAGAGT  | ATAAAGAAAA  | AGTCGGATGC  | 480  |
| 25 | AACCATAAAA | CGTCACTCAA  | ACCTCATATT  | GCTGTGGGAA  | TTAGTGGGAG  | CATTTATAAT  | 540  |
|    | GAAGCAGTAA | TTAAAGAGGC  | ATTCATATT   | GCTCAAAAAG  | AACATGCGAa  | GTTCAGTCTT  | 600  |
|    | ATTTATATAG | ATGTATTTCGA | AAAAAACAGG  | CAATATAAAG  | ATAGTCAAAA  | GCAAGTGCAT  | 660  |
| 30 | CAACATCTCA | TGCTTTGCAAA | ATCATTAGGA  | GCAAAAGTAA  | AAGTAGTTTA  | TAGCCAAACC  | 720  |
|    | GTTCATTAG  | GATTAGACGA  | ATGGTGTAaa  | AATCAAGATG  | TAACCAAATT  | AATTATCGGA  | 780  |
|    | CAACATATTA | GAAATAAGTG  | GCGAGACTTT  | TTCAATACAC  | CTTTAATTGA  | GCATTTAATG  | 840  |
| 35 | TCCTTTGAAC | ATAGCTATAA  | AATCGAAATC  | GTTCCAATCA  | AACAAATACC  | TGTTGAATTG  | 900  |
|    | AAAATGAACA | AATCACCCCTA | TCGTCTTAAa  | GGCAAAcGTT  | TCGCCATAGA  | TATGTTAAAA  | 960  |
| 40 | ATGATTTTGA | TTCAAATAAT  | TTGTGTAATG  | ATGGGACTGT  | GGATTTTATCA | ACTTGATAAG  | 1020 |
|    | CATGAGTCTA | GCTTCATTAT  | TTTAATGATT  | TTTCTCATCG  | GCATCATTTT  | ATTATCCATT  | 1080 |
|    | TGGACGCGGT | CTCCATCATC  | TGGCTTTTAg  | CAGCAATTAa  | TAACGTATTT  | GTgTKTAATT  | 1140 |
| 45 | ATKTTTTCAC | GGAAcCCTA   |             |             |             |             | 1158 |

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 2224 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

## EP 0 786 519 A2

(xi) SEQUENCE DESCRIPTION: SEQ ID No: 485:

|    |  |      |
|----|--|------|
|    | ATCATATGGT CGATTTAACA GATCCAACGT ACTGCLAAAT AATTACATGA CGTTTAAACA  | 60   |
| 5  | TAGCATTGAT TATAACTATT TCTAAGTCTT CGCATTATTT GCGATGATGT GGGAAAGTGT  | 120  |
|    | ATTTTATTTT AAAAATATAA AAAAATAGAT GCAGCAAAAT TTTAAAGCAT TTTATTTTGA  | 180  |
|    | ACATATTAAA AGGGAGCGTA TCATAATGGA ATGTAATGTT TATATCGTAT GCATTACGGA  | 240  |
| 10 | TAAATAATAT ATAAATCATT CTGAGGAGT GAAAGAATAA TGAGAGACTA CACAAAGCAA   | 300  |
|    | TACATTAAAT GCGAATGGGT AGAAAGTAAT AGTAATGAAA CGATAGAAGT TATAAATCCA  | 360  |
|    | GCAACCGAAG AAGTAATCGG GAAAGTTGCT AAAGGTAATA AAGCTGATGT TGATAAAGCC  | 420  |
| 15 | GTCGAGGCGG CAGACGATGT TTATTTAGAG TTCCGTCATA CATCTGTGAA AGAAAGACAA  | 480  |
|    | GGCTTATTAG ATAAAATTGT AAAAGATAT GAAACAGAA AAGACGATAT TGTACAAGCT    | 540  |
|    | ATTACGGATG AATTAGGTGC TCCTTTATCA TTATCTGAGC GTGTCCATTA TCAAAATGGGA | 600  |
| 20 | CTAAACCATT TTGTTCGAGC GAGAGACGCA TTAGATAACT ACGAATTTGA AGAACCCGC   | 660  |
|    | GGAGATGATT TAGTTGTTAA AGAAGCAATC GGTGTATCTG GATTAAATTAC ACCGTGGAAC | 720  |
| 25 | TTCCCTACAA ACCAAACATC ATTAAATTA GCAGCAGCAT TTGCGGCTGG TAGTCCAGTT   | 780  |
|    | GTACTTAAAC CATCTGAAGA AACACCATTT GCAGCTGTTA TTTTAGCTGA GATTTTGTAT  | 840  |
|    | AAAGTCGGTG TTCCTAAAGG TGTATTTAAC CTTGTTAATG GTGATGGTGC TGCTGTTGGG  | 900  |
| 30 | AATCCTTTAT CTGAACATCC TAAAGTACGC ATGATGTCAT TTACAGGATC AGGCCCTACT  | 960  |
|    | GGTTCTAAAA TTATGGAAAA AGCCGCTAAA GATTTTAAAA AGGTATCATT AGAGCTTGCT  | 1020 |
|    | GGCAATCAC CATATATCGT CCTAGATGAC GTAGATATTA AAGAAGCGGC TAAAGCAACA   | 1080 |
| 35 | aCAGGCAAGG TTGTTAATAA TACTGGTCAA GTATGTACAG CTGTACACG TGTTTTAGTG   | 1140 |
|    | CCTAACAAAA TTAAGATGC ATTCTTAGCT GAATTAAGG AACAAATTAG CCAAGTGCCT    | 1200 |
| 40 | GTCGGTAATC CAAGAGAAGA TGGTACACAA GTAGGCCCTA TCATTAGTAA AAAACAATTT  | 1260 |
|    | GATCAAGTAC AAAATTATAT TAATAAGGTT ATTGAAGAAG GTGCTGAATT ATTTTATGGT  | 1320 |
|    | GGTCCTGGTA AACCGAAGG ACTTGAAAAA GGATACTTTG CACGTCCGAC AAITTTTATT   | 1380 |
| 45 | AATGTAGATA ATCAATGAC GATAGCACAA GAWGAAATTT TTGGCCAGT AATGTGAGTT    | 1440 |
|    | ATCACTTATA ACGATTTAGA TGAAGCGATT CAAATTGCAA ATGATACAAA ATATGGTTTG  | 1500 |
|    | GCAGGATATG TTATTGGTAA GGACAAAGAA ACATTGCATA AAGTAGCTCG TCTATTGAA   | 1560 |
| 50 | GCAGGTACAG TAGAATAAAA CGAAGCAGGT AGAAGCCAG ATTTACCATT TGTTGGCTAT   | 1620 |
|    | AAACAATCTG GTTTAGGTCTG TGAATGGGCG GATTATGGTA TTGAAGAGTT CTTAGAAGTG | 1680 |

55

EP 0 786 519 A2

|    |   |      |
|----|---|------|
|    | AGTGCACATG ACTAATTAAG TTTTGTGTAC TGTTTTAATT TTGCAATTTT TATAAATAGA | 1800 |
|    | TTTTGTAATT AAAATAAAAA TTGTCTATAG TTATTCATGT ATTTAAAAGG TTGGGGATTA | 1860 |
| 5  | GCATAATGGG ATTTGTCTAG CACAGTTATT TATGCATTGT CATGCCTATC TATTACTTAC | 1920 |
|    | TAACTAAAAA ATAATGAAAT GGGTGTAAC TATATGCCGT AAAGAGAACG TACATCTCCT  | 1980 |
|    | CAGTATGAAT CATTCCACGA ATTGTACAAG AACTATACTA CCAAGGAACT CACTCAAAAA | 2040 |
| 10 | GCTAAAACTC TTAAGTTGAC GAACTATAGT AAATTAATn AAAAGAAGT TGTCTAGCT    | 2100 |
|    | ATTATGGAAG CACAAATGGA AAAAGATGGT AACTATTATA TGGAAAGTAT CTTAGATGAT | 2160 |
|    | ATACAACCAG ATGGTTATGG TTTTAAAGA ACAGTGAAGT ATTCTAAAGG GAAAAAAGAT  | 2220 |
| 15 | ATTT  | 2224 |

(2) INFORMATION FOR SEQ ID NO: 486:

|    |                               |
|----|-------------------------------|
|    | (i) SEQUENCE CHARACTERISTICS: |
| 20 | (A) LENGTH: 1690 base pairs   |
|    | (B) TYPE: nucleic acid        |
|    | (C) STRANDEDNESS: double      |
|    | (D) TOPOLOGY: linear          |

|    |  |
|----|--|
| 25 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 486: |
|----|--|

|    |   |     |
|----|---|-----|
|    | ACATTACnTT GAATCGAAG TTTCATAAAC GACTTGAATG CCAGTTTCTT TTTCAAAATT  | 60  |
| 30 | CTTAATTAAC TCTGGATCAA TATATTGCC CCAATTGTAT ACGTAAATTT TTGATTTGT   | 120 |
|    | ATGCACTTGT TCTTTAGATT TAAACCAATG ACTTAAAGTA AGACAAAGCA TACCCACAAC | 180 |
|    | TAATGCACCT ATAATGAGTT GTAAAAATCG TTTTATTATT TTACACCTCG CTGATTAGT  | 240 |
| 35 | TTTTTCTTAT TTATCAGGTA TTGAATCAA TAATATCCTA GTATTCTCTAA TACAATAACA | 300 |
|    | GCAAAACAATA ATGTTGAAAT CGCATTAATT TCCATACTAA TTCTTTTCT CGCCATAGCA | 360 |
|    | TAAACTTCAA CTGATAACAC ACTAAAGCCA TTACCAAGTA CGAAGAACT TACTGTGAAA  | 420 |
| 40 | TCGTCTAGTG AATAAGTTAA AGCCATAAAG AATCTCTCTA TAATAGAAGG TAAATATTA  | 480 |
|    | GGAATAATAA TGTTGCTTAA TAATTGTGGT TCAGTCGCTC CTAATCTCT TGCAGCATTT  | 540 |
|    | AACATATTAT TATTATyTC ATACAGTTGT GGTAAAGACGA TAATCACACA TATAGGTATG | 600 |
| 45 | CAAAATGCCA TATGAGATAT TAGAACTGTC CAAAACCTA AACCAAGACC AGTAAATGG   | 660 |
|    | CCAATCGTTG TAAACATAAT TAAGATGAT GCACCTATGA CAACGTGGA TGATACCATC   | 720 |
| 50 | AAGACATTAT TCAATGTTAG TAAAGTACT TTAACCTTTT TATTCTTAA ATAATAATA    | 780 |
|    | GCAATGGCAC CAAATGTACC AATAACTGTA GAAATTGAGG CTGCTAAAG TGCTACAGCT  | 840 |

55



EP 0 786 519 A2

AATGTAAAAT GTTCAAAGTG AATCATATTA CCAGCCGAAT TGAATGAATA GAACATTAAA 960  
AAGAATATTG GGATGTATAA AATCGCTAAA AGTATCCCGA TATACAGCTT TCCATACCAT 1020  
5 TCATATGAT TCACCCCTCTC CCATTAGATG ATTTTGAAT GATTAAAATG AATGCCATAA 1080  
ATACAATTAA GAATATAGCT ATAGTTGATC CCATACCATA ATTTTGAATT GTTAAAAAAT 1140  
GTTCTCTAT TGCCTACCT ATATTTATGA CTTTATACC TGCAATTAAT CTTGTAATCA 1200  
10 TAAATAATGA AAGTGATGGA ATAAAGGTTA CTTGAATCCC AGTCATAACA CCTCTTTTG 1260  
TTAACGGCAT GATTACTTTT CTAAAAGTAT AGAAGGACT GGCACCTAAA TCACCTGAGG 1320  
CCTGCAATAA ATTATTAGGA ATTGCTTTCA TGCTATTAAA TATAGTAAA ATCATAAATG 1380  
15 GTATATAAAT GTAACCTGCC ACTACTAAAA ACGCACCAGT TGTAATAAAC AAATTTGAATG 1440  
ATGGTAAATT AAATAAGTGG AAAATTGATT AATCAGCCCA TCATGACTTA ATAAACCTAT 1500  
20 AAAAGCATAT GTCTTTAACA ATAAATTAT CCATGTTGGA ATAATCATT TCAATTATAA 1560  
GATATTTTGA AATTTCGAAC GAGTAATATA ATAGGCAGnT GGATAACTGA TAGTCAAGGT 1620  
AATAATTGTT ATTGAAGCGG CATATAAAAT TGAATATGCA AACATTTTCA AATATTTTGT 1680  
25 AGTAAAAAAT 1690

(2) INFORMATION FOR SEQ ID NO: 487:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 2112 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 487:

ACGAAAAGAA ATATTATGAT GAACAAAAG AAGAATAAC GATTATATG AAGTACAATG 60  
TGAAGAGTTA TAAAAATATA AGCTTCGCTA ATTTTAAAGA AAACCCAATG GATGGTTATT 120  
40 CTATTAGTGG TTATATAAAT AatGataAAA AGTTATCATT TACAGCTGGT ATAAGATCTG 180  
TTGATGATT TCAATTTGAT ACCGATATTT CTATACAGA TGAATTTGGT AGAAAATTTA 240  
ATAAAAATCC TAAGTCAGTT TCTGAAATAA AAAAAGACA AAATACGTCC AATAAATAAT 300  
45 TGTTCATATT GTGATGAAAC AAAAATATAA GTCATAGAT GAGTTTAACT ATGTTATATA 360  
TATTTGTAGT ATCTATAAAA ATCTCGACAC TATTAAAAATG ATAAAGTGCC GAGGTTTTCT 420  
TACTTATTTA GTTAATTCAA AGTTTATGCC AGATTCATAA GAATTTGTGA CACTTTTAAT 480  
50 AGTGTAACAT TGATTATTAC AATTTATCAA ATGCTCCTTT AGAAGGTATA AATAACAAAA 540

55

|    |  |      |
|----|--|------|
|    | TATTTATTATG TTCAAAACTT TACGCTCCAA AAAGTAAAAA GGAAGTTAAG CAATGTTTAG | 660  |
|    | TTGCTTAaCT TCGGaTATTG AACGCATCAG TCCAATTGGA CATAGAGCCT TTTTATAGTTC | 720  |
| 5  | TTGATGTTTC TCTTTAAAAAC CTTCATATT TTACAAAAAG AAAAATTAGC AGTATAATTA  | 780  |
|    | AGACAACGAA AATAAGTATT TACTTATACA CCAATCCCCCT CACTATTTCG GGTAGTGAGG | 840  |
|    | GGATTTTAT TGGTGGCGCT ATATGTCACC TATTTTGTAT TGGCTCTACT TAGCCAATAA   | 900  |
| 10 | GGAAAAAACG CAATGGCACA GCCACTGATG ACTGGTGCTA TGATGTGAAC GAAATAAGC   | 960  |
|    | ATCACCTTAT ACACCTCCTC TCTGCGTCTA AATTGACGgC TGAGAgGTAG GcGACTCTAC  | 1020 |
|    | TATTATATCA TCGGCAATA TACAAGCACA GTCACITGCT TCTGATAAGT TATATGATTC   | 1080 |
| 15 | TAGCTGATAG ATTGAATCGT CTACACTTAA TTGGACAAAT TCTATGAGAA TAGATATTGT  | 1140 |
|    | TAATTTAAGA AAGTAGGCGA TTTTATTATG ACAAGAGAAA GAAGATCATT TAGTTCAGAG  | 1200 |
| 20 | TTTAAGTTAC AATGGTTAG ATTATATAAA AATGGTAAGC CTAGGAATGA AATTATACGC   | 1260 |
|    | GAGTATGATT TCACACCTTC GACGTTTGT AATGGCGGTT ATAAAAATGTA GGAAATGGA   | 1320 |
|    | TAAAGCAACA TCAAAACACG GGTACATTCA ATCACCAGA TAACCTTACG GATGAAGAAA   | 1380 |
| 25 | AAGAGCTGAT TAAATTACGC AAAGAAGTTC AACATTTAAA AATGGAGAAC GATATTTTAA  | 1440 |
|    | AGCAAGTAGC GCTGATTATG GGGCAAAAAT AGAAGTCATT CAAAAGAATG CACATCAATA  | 1500 |
|    | TTCAAGTATCA GCAATGTGTA AAGTCTGAT AATACTAAGA AGTACCTATT ATGATCTTAT  | 1560 |
| 30 | AAAAAGAAAA GATAATAAAA TCACTAAAGA TGATTCAAAC ATAGAACATG CCGTCATAAA  | 1620 |
|    | TATTTTAAAT TCTAATAGAA AAGTCTTTGG TACAAGACGA ATTAAAAATC ATTTAAATGA  | 1680 |
|    | CAAGGGTCTC ACTGTATCTG GACAAAAGAT AGGTCGATCA TGAAAAAATC TAGTTTCTGT  | 1740 |
| 35 | TTATACGAAA TCTAATACA AAAATCATCT AAAAGAACT AATGAAAAAC GAATTAATAA    | 1800 |
|    | TCTTTATTAT TAGCTGCTGG TGTATTATT GTTAGTCCAA TTTCATTATC TTATAATTCA   | 1860 |
|    | GATGTAGTCT ATGCTGAAGA TAAGTTAGAC CATTCTCAAG CAAAGGTAAT ATATTTGAGT  | 1920 |
| 40 | AACCAAAATT TATTTGATGA ACTTGAGAAA AAAGGTTATA AACTGGAAGA TATATTTACA  | 1980 |
|    | AAAGAAGAAA TAAAAAATA TAAAGCTGAA GACCAATTGA GAGCGGGTAA AACTCAATAT   | 2040 |
| 45 | GTAGAAACAG GTAAAGATAC TGCAACATTA TATCTTCTT CTGCATATAA AAAACAATA    | 2100 |
|    | GCTGCTTAG GT   | 2112 |

(2) INFORMATION FOR SEQ ID NO: 488:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 454 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 488:

5 GTAGGCACAC ATCTGCCATA TAAACATTCT TTTATAC TAGTCTCAT TAGTGTAGAG 60  
 TTATAGTCTC CTCTTGAAT CTCGAATAAT TCAATCAACC TATCAACCTT AGTCTCTTCC 120  
 GTTACTTCTT TTTCAATATC AACTATGAAG GGGATATCAA TTGGAATAAA ACTTGACGTC 180  
 10 GAACACTTAT TTGTATTGG ATGAAAACGA ACGAATCCAT CACTAAATCC TGTGTAAAAA 240  
 AATATTTTTC CTGTGATAG ATCCGGATT TCTCGGCCCT ATTTAATTAA TTCATCTAAT 300  
 CTCATTCTT TTTTAACTTT GATTTCATT GTTATATCTC CTCTTGAACA GTAAATTTAT 360  
 15 CGTTAATGA TACGTATCCA GTCACATTAC ATAAGATGCT ATCAACATCA AAAGTCACAC 420  
 AACAGTTGCG TTCAACATCA TTTGAATAGA ATCT 454

## (2) INFORMATION FOR SEQ ID NO: 489:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1372 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 489:

30 TTGTCAGAAAT TAGAATGCTT TTGAGTTACT TCATAATACT CATCAGTTT TTGTGTATCC 60  
 TTTTGACTTT TATTTATTC TTTCCACTTA CCGATATGAC TTTCTTTTT TACAGTTATT 120  
 TTCGTTTGT TTAATAAATG ATAACCGATG TTCCTTTTGG TATCGTGGGA CTCTAAAAAG 180  
 35 ACTGAATTGT TTTCTGATT ATCAGAAATG GTTGTTTGTT TATCGTCTGT ATATAACGTA 240  
 TACCCATTGG CTITGGGATT TTCAATCGTT GTAACGGAT TCTTTGATGA ATCAGTACTT 300  
 TTAATGCCAG TTCTAAGAA GACAATTTTA TCGTTTAATA TGAATATGA TTTTITGGCA 360  
 40 GTTAAAGTTT TGCTCTGATT TTCAAAATCC ATTCGGATAC TAGCATGTTG GTCATCAACT 420  
 TTTGTTCCGC CAACAAAAGT TTTACTCGAC TTTTATCAT CCGTATCTTT TAATATTTCA 480  
 TTGTTCTAAG TTGTTGTACC TGATAAACGT TTCATATCGG CTGTCACCCA GAAGTTATCA 540  
 45 TGATAGTGTT TGACATCGCT GTTATATAAA TAAGACATTC CAGCACCAGT GTGCCAACCT 600  
 TTTAAATCTC CTCGGTTGAT ACTTTCATAG CGTGCTACGT TTTTCGAGT CATACTTAAA 660  
 CCAATGCGAA AGTCTAAGTC TTTGTTATGA TAGGTGACAC GATCCATGTC ATTATATATT 720  
 50 TTAAGTTGTT GTGTTAATCC GTTTTATGAA ATACTGTTAT CTGTCATTAA AGACTTCATT 780



EP 0 786 519 A2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 491:

|    |            |             |             |             |            |            |      |
|----|------------|-------------|-------------|-------------|------------|------------|------|
|    | TATCCACCCC | ACGAGAGCCC  | CGGAAACTTA  | TTGTGTTACA  | AGATATATAA | GCAGAAACGA | 60   |
| 5  | ACAACAGTTA | ACAAAATAAA  | TGAAATTA    | CGTTTTAAAA  | ATGAAACAAA | TGAAATCATC | 120  |
|    | TATTAGGTTA | TGAACTGTT   | TATAGCTTGA  | ATAGAAGCAT  | TTATTTTTTA | GGAGGACAA  | 180  |
|    | TATTATGCGT | CAAAACATTA  | TGGCAAATGA  | ATCAAAACAT  | GAGCGCAAT  | GGTATGTTAT | 240  |
| 10 | CGATGCTGAA | GGCCAAACAT  | TAGTTCGTT   | ATCATCAGAA  | GTAGCATCTA | TCTTACGGCG | 300  |
|    | TAAAAATAAA | GTAACCTTACA | CACCACACGT  | TGATACTGGT  | GATTATGTAA | TCGTTATTAA | 360  |
|    | TGCATCAAAA | ATCGAATTTA  | CTGGTAACAA  | AGAAACTGAC  | AAAGTTTACT | ACCGTCACTC | 420  |
| 15 | AAATCACCCA | GGTGGTATCA  | AATCAATCAC  | TGCTGGTGAA  | TTAAGAAGAA | CTAACCAGAA | 480  |
|    | ACGTTTAATT | GAAAACTCAA  | TTAAAGGTAT  | GTTACCAAGC  | ACTCGTTTAC | GCGAAAAACA | 540  |
|    | AGGTAAAAAA | TTATTTGTAT  | ATGGTGGCGC  | TGAACATCCA  | CACGCTGCAC | AACAACCAGA | 600  |
| 20 | AAACTACGAA | TTACGTGGTT  | AATTAGAAGG  | AGGAAATGAC  | TTTGGCACA  | GTGAATATA  | 660  |
|    | GAGGCACAGG | CCGTCGTAAA  | AACCTCAGTAG | CACGCTGACG  | TTTGTACCA  | GGTGAAGGTA | 720  |
|    | ACATCACAGT | TAATAACCGT  | GACGTACGCG  | AATACTTACC  | ATTGGAATCA | TTAATTTTAC | 780  |
| 25 | ACTTAAACCA | ACCATTTGAT  | GTAACGTAAA  | CTAAAGGTAA  | CTATGATGTT | TTAGTTAACG | 840  |
|    | TTCATGTGGT | TGTTTCACTG  | GACAAGCTCA  | AGCTATCCGT  | CACGGAATCG | CTCGTGCAAT | 900  |
| 30 | ATTAGAAGCA | GATCCTGAAT  | ACAGAGGTTT  | TTTAAACGCG  | GCTGGATTAC | TTACTCGTGA | 960  |
|    | CCCACGTATG | AAGAACATA   | AAAAACCAGG  | TCTTAAAGCA  | GCTCGTCGTT | CACCTCAATT | 1020 |
|    | CTCAAAACGT | TAATTGTCGG  | ACGATATATA  | CAAAACACCT  | CGATATTATG | TCGAGGTGTT | 1080 |
| 35 | TTTTTGGCGT | TTTTGCGGCG  | AATATGGAAT  | GTGTAGAATA  | TAAATGAATT | TTTACCTTCC | 1140 |
|    | CACCATAAAA | GATGAAGAAC  | CATGAATGTG  | GAGAAACAATA | AATAGTTGGA | TATTCTGTTA | 1200 |
|    | TTTTTTTGGG | AGTGAAGTGG  | GATTTGGAAT  | ACTTTACTCA  | AAACGATTAA | AAGGTTTAAA | 1260 |
| 40 | AAAAACAACA | AAAGAAA     |             |             |            |            | 1277 |

(2) INFORMATION FOR SEQ ID NO: 492:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 673 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 492:

TTATTGACAT TGTTTTATC CAAAATTCAT TGTTAAGACA TTTTCTTTAT GAAATAATAA 120  
 TtATTATTGAA GTATATTTTT ATTATTATTA AAAATAAATA AGGGGATACT TATGAGCACA 180  
 5 AATCAAAAT TTTTAAATATT TGTTATAGCA ATTATCTAC TTACATCTGT AATAGGAATT 240  
 GTTGGAGCAT ACATGAGTCG TCAACGTCTA TTAATACTA TGGAAACATT ATGGCAAACG 300  
 ATTCTCCAT TAGAAGCTTT TATCAGACCG AACTCACATT TCGACTATGA GTATAAGCTC 360  
 10 TACAAGGAAA AATTGAATC ACATTCATTA GTTGATGATA AAACCTGGTC CGACTTAAAT 420  
 ATGAATGCAA TCTTTCATAA GATGAATTAT AATTAAACAG CTATTGGTGA AATGAAGCTA 480  
 TATGCTGTT TACGTGGAAT GCTTCAATT ACGAACAAT CATTACTTAG TTTATTTAAT 540  
 15 GATAATGCTG AATTAGAAA AAACGTAACA TATCATTTAG CTTTGATTGG TAAAACGTGT 600  
 ATCCAACATT TCCAGACCAA ATCACACCGG KAAACGTCCA AATATAATTG TCTATGCCCG 660  
 GTTACCAGT ATC 673  
 20

(2) INFORMATION FOR SEQ ID NO: 493:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1240 base pairs  
 (B) TYPE: nucleic acid  
 25 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 493:

30 AAAAAAAGTT AATGCTGACG GTGTATTAAC TTTTGATATT CTAGAAAATA AATATACTTA 60  
 TGAAAGTTATT AACGCTATAG GGAAGAAGATG GATTGTTAGT CATGTCGAAG GTGAAAACGA 120  
 35 CAAGAAAGAA TATGTAATAA CTGTCATTGA TAGGAAATCA GAAGGCGACA GACAACTGGT 180  
 TGAATGTACT GCTAGAGAGA TTCCCATAGA CAAGTTAATG ATTGATAGAA TTTATGTTAA 240  
 TGTAAACAGGA TCTTTTACAG TAGAAAGATA TTTTAACATT GTGTTTCAAG GTACTGGAAT 300  
 40 GCTTTTGTAA GTGAGGGGCA AAGTTAAATC TTCAAAGTTT GAAAATGGTG GTGAAGGCGA 360  
 TACAAGGTTA GAAATGTTTA AAAGGGGATT AGAACATTTC GGTTTAGAAT ATAAAATAAC 420  
 GTATGACAAA AAGAAAGACA GATATAAGTT TGTATTGACG CCTTTTGCAA ATCAAAAAGC 480  
 45 GTCTTATTTT ATTCTGACG AATCAACGCC AACGCTATAA AACTCGAGGA AGATGCAAGT 540  
 GATTTGCCCA CCTTCATTAG AGGATATGGT AATTATTCAG GAGAAGAAAC ATTCGAACAC 600  
 GCTGGGCTCG TAATGGAAGC TAGAAGTGCA TTAGCTGAAA TATACGCGCA CATCCACGCA 660  
 50 GAACCATTTA AAGATGGTAA AGTGACTGAC CAAGAAACTA TGGATAAAGA ATTACAATCG 720

55

|    |             |            |            |             |             |             |      |
|----|-------------|------------|------------|-------------|-------------|-------------|------|
|    | TATCCGAGAA  | CAGACCACA  | ACCCGAGAC  | ATAGTACAAA  | TAAATCTAC   | CAAACTAGGT  | 840  |
|    | TTGAATGTATT | TAGTCGGTAT | AGTACAGGTT | AAAACGATTA  | GGGGTATAAA  | CAATGTAATT  | 900  |
| 5  | GTTAAGCAAG  | ATGTAACGCT | TGGTGAGTTT | AATCGAGAAC  | AACGATTATAT | GA AAAAAGTT | 960  |
|    | AATACTGCAG  | CTAACTATGT | TTCTGGATTA | AATGATGTTA  | ACCTTTCTAA  | TCCTAGTAAA  | 1020 |
|    | CGCGCAGAAA  | ACTTGAAGTC | TAAAGTAGCG | TCAATAGCTA  | AATCAACACT  | CGATTTGATG  | 1080 |
|    | AGTGAAACTG  | ATTTGTATTG | AGATAAACAA | CAGAAAGGTA  | GCTCTAAAAAC | TGTGACTACA  | 1140 |
| 10 | TCTGACGGCA  | CTATCGTTCA | TGATTTTATA | GATaAATCm   | ACATTAAaG   | TGTAAAmCG   | 1200 |
|    | aTTGGAAACga | TGGCGGATtC | TGTAGCTAGA | GGATCAACATG |             |             | 1240 |

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1311 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

|    |            |            |             |            |             |             |     |
|----|------------|------------|-------------|------------|-------------|-------------|-----|
|    | ACGGTGGATT | TAGCCTAGAA | GATTTAACGC  | ATCAAGGTaa | ATTATCAGCA  | TTTAGCTTTA  | 60  |
|    | ATGATCAAAc | AGGTCAAGCA | ACATTGATTA  | CTAATGAAGA | TGAAAACCTC  | GTAAAAGATG  | 120 |
| 30 | AGCAAGCTGC | TGGCGTAGAT | GCAAAATTATT | ACGCTAAACA | AACATATGAT  | TATTACAAG   | 180 |
|    | ACACATTTGG | TCGTGAATCA | TATgaCAACC  | AAGGTAGTCC | AATTGTTTCA  | TTAAGCGATG  | 240 |
|    | TTAATAACTA | CGGTGGTCAA | GATAACAGAA  | ATAATGCCGC | ATGGATCGGT  | GACAAAAATGA | 300 |
| 35 | TCTATGGTGA | TGGTGATGGT | CGCACAATTCA | CAAGTTTATC | GGGTGCCAAAT | GACGTAGTAGT | 360 |
|    | CACACGAATT | aACACACGGT | GTTCACAGA   | AGACCGSGAA | CTTAGAATAT  | AAGGACAGAT  | 420 |
|    | CAGGCGCTCT | AAATGAAAGC | TTTTCAGATG  | TTTTTGATA  | CTTTGTAGAT  | GACGAGGATT  | 480 |
| 40 | TCTTAATGGG | TGAAGATGTC | TACACACCTG  | GAAAAGAGGG | AGACGCTTTA  | CGCAGCATGT  | 540 |
|    | CAAAACCAGA | ACAAATTGGT | CAACCAGCTC  | ATATGAAAGA | CTATGTATTTC | ACTGAAAAAG  | 600 |
|    | ATAATGGTGG | CGTACATACG | AATTCTGGAA  | TTCCAAATAA | AGCAGCTTAT  | AACGTGATTG  | 660 |
| 45 | AAGCAATAGG | GAATCTTAA  | TCAGAACAAA  | TTTACTACG  | AGCATTAAAG  | GAATACTTAA  | 720 |
|    | CAAGTAATTC | AAACTCTAAA | GATTGTAAAG  | ATGCATTATA | CCAAGCGGCT  | AAAGATTGAT  | 780 |
|    | ATGACgAGCA | AACAGCTGAA | CAAGTGTATG  | ATGCATGGAA | TGAAGTGGCG  | GTGGGATTAA  | 840 |
| 50 | AATATATAAA | CaAGAGGAAG | TAGTGTAAAA  | CACCTATAAA | TAATTAAAGT  | TTAAATACAG  | 900 |

10  
 15  
 20  
 25  
 30  
 35  
 40  
 45  
 50  
 55

ATTAGATGAG AGGAGTGTGA GGGTGTCTG CGAAAGACT ACTCGGCAGT CTAATCAT 1020  
 TACAAGTAGT AGATATGTGA TAATTAAATG CTGACTTAGA ATACAAAATT CATTTTAAAA 1080  
 GTTGTCACAA AAAATTTACA TGTATTTTTA TTATCTTTTG CAAAACAAAG TGTAAATTA 1140  
 TAAATGAAC ATGCATGAAT TTATTTTTTA ATACAAGAAA CGTAATACC AAAGGAGTTT 1200  
 ACAATATGAA GAAAAGTAAA CGATTAGAAA TTGTTCTTAC AATAGTTAAA AAGCATAAGA 1260  
 TTTATAAAAA AGAACAAATc ATTTCATATA TTGAAGAATA TTTTGGTGTA A 1311

(2) INFORMATION FOR SEQ ID NO: 495:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1761 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 495:

10  
 15  
 20  
 25  
 30  
 35  
 40  
 45  
 50  
 55

TGCACCTTCT AAAATACTT GCCTTACTTG TTCCAATCC TTGTCAGACA TTGATTCGGC 60  
 TAATTTATTC ATTATGCCT CAAATACGTT AATTATGTCA TCCATTTTCAG TACTATAAAA 120  
 ATCAAACCAT TTGCGGTAT CTTTTCTCT ATTAAAGTTA TGATCACTTT GAGAGCGCTT 180  
 AGCTAATCTC GCATAAATAT ATGGACAAGG TGCCATTGCA GCAATTGTAT AAATAGCATT 240  
 TTCACGACTA TGCCTTGGAA AATACATATG TTTTATGTAA TGGTCGCCAC TTGAGGAGCCA 300  
 AACTTTTGTG TTAATGATTT CTTCGTATGA TTCACCAACA ATTTGCGCTA AAATATCATG 360  
 CGCAAGTACT TCACCTTCAA CCATAAATTC TATTTGCTCT ACTAAAAAT TTAGCTCATT 420  
 CATGCTATTC ATTTTGGAA TTAACAAAGC ATATAAGTTT GTAAATTCTT TTAATACGC 480  
 AGCATCAGCT TTTAAGTAAT GCGTAATGC GTCAGCTCCT ATATCTCCG ATAACATCTT 540  
 CTGAATAAAG TCATCTCAT AAATATCATT AATGATTGGC TTGTCAGCTT GGTACAATTT 600  
 TTGTGAAAAA TCCATTGTAA AAAATCCTCC CTAATAAAA AAACACTTCC CAACATGAAA 660  
 GTAGTTTGAT GGCAATGTTG CTATACTAGC CCCATCACTT CAATACTAC TTTCCTACGT 720  
 TGGTACTAAC CAAATCAGGT CATAAGGGTC TGAACAATTC ATCTCAGCCA TATCATTAGG 780  
 CTCGCCTAGT AGTTCCTTAG TATTCATTTG CAAATTAATC TTAGCAACAG GTTCAACAT 840  
 TTTCAATTAT TGTGCTCAG TTGTATTATT ATCTTTAAAT AATAATTCTA TAATGACATA 900  
 TATTGCGAA ATAAAAAAC CGAACATAT CGAGAATCC CCGATATATT CCAATCTAAA 960  
 AGTACTTAT ATAATATTA ATAGCTATG CATAAATGGC TTATGCAGTA ACCCAATGTC 1020



|    |            |             |             |            |            |             |     |
|----|------------|-------------|-------------|------------|------------|-------------|-----|
|    | TCATTTATGA | AAATATGTCGG | ANAGAnCcaA  | GaAmaCAAT  | TGAGCGTGAA | GA AAAAGCAA | 60  |
| 35 | GACTTAAAGA | AGAACAAAAG  | GCACGTCAAA  | ATGAACAGCC | ACAAATAAAA | GATGTGAGTG  | 120 |
|    | ATTTTACGGA | AGTGCTCTCA  | GAAGAAGATA  | TTCCAAATTA | TGGGCATACT | GAAATGAAA   | 180 |
|    | GTAAAGCCA  | GAGTCAACCA  | AGTCGAAAAA  | AACGAGTGT  | TGATGCAGAG | AATAGTTCGA  | 240 |
| 40 | ATAACATCGT | AAATCATCAT  | CAAGCAGATC  | AGCAAGAACA | ATTAAACGAA | CAAACCTATA  | 300 |
|    | ACAGTGTGTA | AAGTGAAAAA  | ACTATTGAAG  | AAGCTGGTGA | AGTTACGAAT | GTATCGTATG  | 360 |
|    | TGTGTCCACC | GTTAACTTTA  | CTTAATCAAC  | CTGCAAAAAA | AAAAGCAACA | TCTAAAGCTG  | 420 |
| 45 | AAGTGCAACG | TAAAGGACAA  | GTACTAGAGA  | ATACNTTAAA | AGATTTTGGG | GTAATGCAA   | 480 |
|    | AAGTGACACA | ATTTAAAAAT  | GGTCTGCGAG  | TAACTCTAAT | TGCAATTCAG | CCAGCTCAAG  | 540 |
|    | GGGTAAAGT  | GAGTAAATAT  | GTTAACTCTG  | ATAATGATAT | TGCATTAGCT | TTAGCAGCAA  | 600 |
| 50 | AAGATGTTAT | AATCGAAGCG  | CCAACTACCTG | GTGTGTCCTG | AGTAGGTATT | GAGTGGCCAA  | 660 |

# EP 0 786 519 A2

ATAAATAGAG AGTTGGATTG GGAAGAGATA TATCAGGTGA TCCAATTACT GTTCCACTAA 780  
ATGAAATGCC ACAC 794

5 (2) INFORMATION FOR SEQ ID NO: 497:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1161 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
10 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 497:

15 AGCCAGTTTT GcATTGTC AAATCGCAAT GAATATATTG ATTGCTTGTC AAAATTGGGA 60  
AGAAATTAT TCTCTTTTT CTTTTAGCA GTTATCATTC CTAATTGTC TTTCTTTTGA 120  
TCTTAGTAT AAGAATCCA TCTGAATTAA AGTCTGCGG CATGAATTGT AACGTTTTGA 180  
20 CCAACTCTCC AGTTATCGGA TGTGAAACG GTTCAAAATC GAAGTTTTTA TTAATTTTCA 240  
AAAACGTATA AATCAGGTTT TCATTTTCTA GTTGCTCAAT TGTACATGTT GAATAGATGA 300  
TTTCTCCACC TATTTTACA TTGTTTTTGA CATTTTCAA TATTTCAAGC TGTAAATCAA 360  
25 CTAGTACTC AATATGTTGT TTGCTTTGAG TATACTTAAT CTCGCGCTTA TGTCCTAATA 420  
CACCTAATCC GCTACATGTT GCATCAACAA GTATCTTATC GTATGTTTTA TCATAAGGTT 480  
TTGTGCGATC ATGTTGAAAA GCTTTAATAT TTGTTAATCG TAATTTTAT ATATTAATAA 540  
TAATTAAGTC TATTTTGTTA TCATGTATAT CTGAAGCGTC AACITGCCCT TCTGGCATT 600  
AAACTTCAGC AATGTGACAA GCTTTACGCG CAGGTGCAAT ACATGCATCT AATACGTGAT 660  
35 CATGTCGGTC TACATTCATA ATGTGTGCAa CAAACATTGA GCTTTTATCT TGAATTGAAA 720  
CGAATCCATC TTAAATGAA CGAGAATGAA TAATTGGTTG TCCTCCTATA TGGAGACAAT 780  
AAGGTAAGTC ATGATCTTTT TCAACGTCAT AACCTTCGTC TTGCAACTTT TCAATAATAT 840  
40 CATCTAATGA TGCTCGCGTC AGGTTGGCAC GCACAGTTGT TGATGTGTTT TCTAAAAATG 900  
ACTGTAAAAA TTTTTCAGTT TCTTCGAGAC CATAATGTTG TGCCCAATGA TCTATAATCC 960  
ACTTCGGCAT ACTACTACG ATTGCCATTC TTTTITTAGG ATCTGCAATT TCAATTAATA 1020  
45 CAGGTAAGTC ACTACGCATC ATTGTACGTA AAATACCATT TACGACATTA CCATTATGAT 1080  
AGCCACCGCG TTTCTTTGCT ATTTCAACTG CTTCATTAAT AATGGCATGA TTTGGAACCT 1140  
TATCTAATAA TACATATTGA T 1161

50 (2) INFORMATION FOR SEQ ID NO: 498:

55

# EP 0 786 519 A2

(A) LENGTH: 1504 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 498:

|    |  |      |
|----|--|------|
| 10 | AGCTCAGCTC ATCTTCGGCG GCGCTAAATT AAAATAATCA ATTTCTGAGT TAAACTTTTA  | 60   |
|    | TTTACAACAT ACTATTACTA TACATTACAA ATTTTAAAAA TATACATTAC ACTCATTACT  | 120  |
|    | CAATGGAAG CGTATGATTT CACAGCCCCC CTAGCTTGTA GAAATCATAC TTTCTTTT     | 180  |
| 15 | TCAATATATA TACAACTATT AAATCCCAT AAGATTGCAGA GCACATAAGT AAATTTTTTT  | 240  |
|    | AGAGCTTGAG GTTTGTTTAG CTTAAGCAAC CCATGAGCTC AACACTTCC TGTTACATA    | 300  |
|    | ACACTACAAA TCGCATTATG TTGCTTAATC TTATGTTTAT ATAAATTACA CACAATAAAT  | 360  |
| 20 | AGAAAGAATG TGAACATCAT GAATAAAITA TTGCTACTCG TTACATTTAT CATTCGTGTG  | 420  |
|    | GGTTCAGGTA TTGTTATGTT AATGCAAGGC TACGAAAAAT TAACGGGCGG ATTACGCTG   | 480  |
|    | AAAGGTTTAG TACCAGTCAT CGCTAACAACT ACTGATTCAC CAGAGTGGTA TAAGTGGTTT | 540  |
| 25 | TTGCAAAATA TAGTTGCACA TACGACGTCA TTATTTGATA TTGTTGTCCT ACTCGGAGAG  | 600  |
|    | ATTGCAATTG GATTAGGTTT AATTTTGGGA GTTTTTCAT ATGCTGCTAG TTTCTTTGGA   | 660  |
| 30 | GCCTTTGTTA TGATAAAITA TATCTTAGCA GATATGATAT TTACGTATCC TCTTCAATTA  | 720  |
|    | ACTTTCTTTA TCTTTTACT AATGAGTCAC TCATTGTAA AACAGATTTC ACTTAAAGAA    | 780  |
|    | ATCATTAATT ACTTTAGAGG TCGTAAGAAC AGAGGTGAAA AAATAGATGA CCCACTTACT  | 840  |
| 35 | GATCGTGGAT GATGAACAAG ACATTGTAGA CATTGTCAA ACCTATTTTG AATATGAAGG   | 900  |
|    | TTACAAAGTA ACAACGACAA CTAGCGGTAA AGAAGCAATT TCTTACTAT CAAATGATAT   | 960  |
|    | TGATATCATG GTACTTGATA TCATGATGCC AGAAGTTAAT GGTTACGACA TTGTCAAAGA  | 1020 |
| 40 | AATGAAAAGG CAAAAATTAG ATATCCCCCTT TATCTATTTA ACTGCCAAAA CACAAGAACA | 1080 |
|    | TGATACCAAT TACGCCTTAA CTTTAGGTGC AGATGACTAT GTCAAAAAAC CATTTAGTCC  | 1140 |
| 45 | AAGGGAACCT GTTTTACGTA TTAATAATTT ACTTACAAGA ATGAAGAAAT ACCATCATCA  | 1200 |
|    | ACCAGTTGAA CAACTGTGCT TTGATGAATT AACACTTATT AACTTAAAGTA AAGTGTGAC  | 1260 |
|    | GTAAATGGT CAAGAAGTCC CTATGCGTAT TAAGGAATTT GAGTTATTGT GGTATTTAGC   | 1320 |
| 50 | TTCTAGAGAA AATGAAGTTA TTCTTAAATC AGAATTACTT GAAAAAGTTT GGGGATATGA  | 1380 |
|    | CTATTACGAA GATGCTAATA CCGTGAATGT CCATATACAC CGTATTAGAG AAAAAATAGA  | 1440 |
|    | AAAAGAGAGC TTTACAACAT ATACCATCAC AACTGTATGG GGATTAGGAT ATAAATnTGA  | 1500 |

55

## (2) INFORMATION FOR SEQ ID NO: 499:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1623 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 499:

ATTGAAAGCG ATAATTCGTA nTAATTGAGT TTGTTGAAAA ATTTAGGGTA ATGTAAAGAT 60  
 ATAAAGATA CATAGAYTGG AGAGATATAA AGATGTTGAA TGAGATACAA ATATTAaATA 120  
 aTGGATACCC GATGCCTTCA GTTGGGTTAG GTGTTTATAA AATCTCTGAC GAAGATATGA 180  
 CTAAGATTGT AAATGCTGCA ATTGACGCAG GCTATAGAGC GTTTGATACA GCATACTTTT 240  
 ATGATAATGA GCCTTCACTA GGACGAGCAT TAAAGGATAA TGGCGTCGAT AGAGAAGATT 300  
 TGTTTATAAC AACGAAGTTA TGAATGACT ATCAAGGTTA TGAGAAAACA TTCGAATATT 360  
 TCAACAAATC GATTGAAAAA TTACAAACTG ATTATCTTGA TTTATTCTTA ATACATTGGC 420  
 CTTGTGAAGC AGATGCTCTA TTTTAGAAA CATATAAAGC TATGGAAGAA CTTTACGAGC 480  
 AAGGTAAGGT AAAAGCAATA GGTGTATGTA ATTTTAATGT TCATCATCTA GAAAAATTAA 540  
 TGGCTCAATC AAGTATCAAA CCAATGGTGA ATCAAATTGA GGTACATCCA TATTTTAAACC 600  
 AACAGAATT ACAAGAATTT TGTGATCGTC ACGATATTAA AGTGACTGCA TGGATGCCTT 660  
 TGATGAGAAA TAGAGGACTA CTAGACGACC CTGTATTGTT TAAAATTGCT GAAAAATATC 720  
 ATAAAACACC AGCACAAAGT GTATTACGTT GGCATTTAGC ACACAATAGA ATTATTATTC 780  
 CAAAAATCTCA GACACCTAAA CGCAITCAAG AAAATATAGA TATTTTAGAT TTTAATTTAG 840  
 AATTAAACGA AGTAGCTGAA ATTGATGCTT TAAATAGAAA TGCAAGACAA GGTAAAAATC 900  
 CAGATAGTGT GAAATTTGGG GATTTAAAAA AACTGGATGT TAAATTTTAC GTTTATGAAT 960  
 GCCTTTTAAAT GTGTACATTA AAATAAATGA GTTGTTTTTT ACTATTGTAT AAAACAATAC 1020  
 TCAGGTACAT TCAAAATCTT TTAAATAAAA AGGATGGACA TAGATGAAAA TTAGAGTCGT 1080  
 CATTCTTTGT TTTAATGAAG GGAAGTCAT TACACAAACA CATCAACAAT TAACTGAAAT 1140  
 ACTTTCACAA GATAGTAGTG TGAAGGGCTA TGATTATAAT ATGCTTTTCA TAGATGATGG 1200  
 TAGTACGGAT ACCACTATAG ATGAAATGCA ACATCTTGCC ACAATAGATA GGCATGTCAG 1260  
 CTTTATTCTT TTTAGTAGAA ATTTTGGAAA AGAAGCAGCT ATGATTGCAG GTTACCAGCA 1320  
 TAGTACTGAA TTTGATGCAG TCATCATGAT AGATTGTGAT TTGCAACATC CACCTGAATA 1380

EP 0 786 519 A2

TAGAAGTGGT GAAAAATTTTA GTCGAAAAC ATTAAGCCAT TTGTATTATA AGTTAGTTAA 1500  
 TTGCTTTGTA GAAGAAGTAC AATTGTATGA TGGTGTGGT GATTTTAGAC TTTTAAGCCA 1560  
 AAGAGCTGTT AAATCCATTG CATCACTGA AGAATATAAT CGnTTTCAA AAnGGnTATT 1620  
 TGA 1623

(2) INFORMATION FOR SEQ ID NO: 500:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 605 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 500:

AAAGTnGGTG AAnCTATATA CTTAATCTAT ATTTATATAT TAACCATTAG GGTTAAAAAT 60  
 TACTCTAGCA TTTATGAATA GATGGGAGTT TATTTTATTA TTATATAGGA GAGATGTTGA 120  
 ATGACACATC GCGCACTATT AGTTGTTGAC TATTCATATG ACTTTATCGC AGACGACGGC 180  
 TTACTAACAT GCGGTAAACC TCGACAAAAT ATTGAAGATT TTATTGTTTC TCGTATCAAT 240  
 GACTTTAATT ATTATCAAGA CCATATATTC TTTTTrTGG ATTACATTA TTTACATGAC 300  
 ATTATCATC CTGAAAGTAA ATTATTCCCA CCACACAATA TCGTAGATAC AAGTGGTAGA 360  
 GAATTATACG GTAAAGTAGG TAAATTATAC GAAACAATTA AAGCGCAACC TAATGTACAT 420  
 TTCATTGATA AAACGGGCTA TGATTCGTTT TTTGGTACCC CGCTTGATAG TTTATTGAGa 480  
 GAAAGAAGTA TTAATCAAGT CGAAATCGTT GGTGTATGTA CCGATATTGG CGTGTACAT 540  
 ACAGCAATTT CTGCATACAA CTTAGGtTAT AAAATTTAG TACCTGCTGA GGGAGTGGCT 600  
 CATTT 605

(2) INFORMATION FOR SEQ ID NO: 501:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1739 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 501:

TAGGTTnAAA GCATAGnTTT nTCAAAAAAGA CAAATCATTC ATATATTGGA GGATATTTTG 60  
 GTGTAAGATA TAGTGCAACC ACAATTGCTA AAGACTTGAA GGAACATAAT ATATATCGTG 120

TGAGAGAAAA GTTTAGACAC TATTGTGAAC ATGAAGTTCT AAGTTCAATC ATCAATGGTT 240  
 CATACATTAT CGTCAAACC TCACCTGGTT TCGCCCAAGS CATAAACTAT TTTATCGATC 300  
 5 AGCTAAATAT AGAAGAGATA TTAGGTACGG TGAGTGGAAA TGACACTACA TTAATCTTAA 360  
 CTGCCTCAAA TGATATGGCA GAATACGTAT ATGCAAAATT ATTTAAATAG ACATGTATCA 420  
 AATGAATAAT AAAAATTGTT TTGTTATCAC GTGTACTCAA GTTAGTTACC AAATATTAAC 480  
 10 TTGTGTACGC GTTTTTTTAT GGAAGAGAAAG AATTTCATAGT CATTCAATTG ACTGTATAAA 540  
 AAACCTTATA CAACATGTTT TTATGGGTAT TTTTGAATAA AAAATGTATA TTTTGACCCA 600  
 AAATACCTTT ATTTATGTAT AAAAATCCAT TATTATGTAT TGTATAACAA AAAGATATGA 660  
 15 AATTTTCGAC TTCTTTATG TGAATATAAT CACATGTAAG CGTTTGAAGA TTGCTATATAC 720  
 TCTAAATGAA TTCAAAGATA AAAGGAGGAA ATAGACATGA CAGATGTGTC AATTAAGTGA 780  
 AATAGCGAAA TTGGAGCTTT AAAAAGTGTG TTACTTAAGC GTCCTGGaAA AGAATTAGAA 840  
 AATTTAGTAC CTGATTATTT AGATGGATTA CTATTGTATG ATATTCCATA TTTAGAAGTA 900  
 GCTCAAAAG AGCATGACCA TTTTGCGCAG GTGCTAAGAG AAGAGGGTGT TGAAGTACTT 960  
 25 TACCTTGAGA AGTTAGCAGC TGAAGTATT GAAAATCCTC AAGTAAGAAG TGAATTTATT 1020  
 GATGATGTAT TAGCAGAGTC TAAAAAACA ATATTAGGTC ATGAAGAAGA AATTAAGGCA 1080  
 TTATTTGCGA CACTTTCTAA TCAAGAAGTT GTAGATAAAA TAATGTCAGG GGTACGTAAG 1140  
 30 GAAGAAATTA ATCCGAAATG TACACATCTA GTAGAGTATA TGGATGATAA GTATCCATTC 1200  
 TATTTAGATC CAATGCCAAA CCTTTATTTT ACTAGAGATC CACAAGCCTC AATGAGCAC 1260  
 GGTATAACAA TCAATCGGAT GTTCTGGAGA GCACGACGAC GAGAATCAAT ATTTATTCAA 1320  
 35 TATATTGTAA AGCATCATCC TAGATTAAAG GATGCGAATA TTCCAATCTG GTTAGATCGA 1380  
 GATTGCCCAT TCAATATTGA AGGCGGGGAT GAACTTGTTT TATCTAAAGA TGTCTTGGCT 1440  
 40 ATAGGCGTTT CAGAACGTAC ATCTGCACAA GCTATTGAAA AGTTAGCGCG ACGTATTTTT 1500  
 GAAAAATCCG AGGCGACGTT TAAAAAGTA GTAGCAATTG AAATTTCAAC TAGTCGAAGT 1560  
 TTTATGCACT TAGATACAGT ATTTACAATG ATAGATTATG ACAAATTTAC AATGCATTCA 1620  
 45 GCCATTTTAA AGGCAGAAGG CAAATATGAAT ATATTATTA TTGAATATGA TGACGTAAT 1680  
 AAAGATATTG CCATCAAAACA ATCTAGTCTT TAAAAAGATA CTTTAGAAGA CGTACTAGG 1739

(2) INFORMATION FOR SEQ ID NO: 502:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1745 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double

|    | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 502:                         |      |
|----|--|------|
| 5  | CTGTACATAC AGCAATATCG TTAACAATG AAAACAGTAT TTTAGGATTG TAAACCATGA   | 60   |
|    | TAAACCTAAA ATACTGTTAT TTTTATTACT TAAATTTCTT CTTCAGTGCC TTTCACACAT  | 120  |
| 10 | AAGTGGAAC GAATTCAGAA ATATCTGCTC GATAAGCTGC AACTTCTTTA ACAATACTTG   | 180  |
|    | AACTTATAAA TGAATAATTA GTACTAGACA TCATATATAA CGTTTCAATT TCATTGTTCA  | 240  |
|    | ACTTTTATT CATTGAAGTT AAGCGTAATT CATATTCAAA ATCATCTAGT GCTCTTAAAC   | 300  |
| 15 | CACGATGAT GTTTTATGCT CCTACTTGTT CACATAATC GACTGAGTAA CCACTAAATT    | 360  |
|    | GATGAACCTT GACATTAGGT AAATGTTTAA CAGATTGGTC AATTAAATCC ATTCCTCTTT  | 420  |
|    | CTAAACTAAA CGTACCTTCT TTTTACTAT TTTTAAAGAC ACAGACATGA ATTTCATCAA   | 480  |
| 20 | ATCTATCTGT ACTTCTCTCA ATAATGTCTA AATGACCATA AGTAATGGGG TCAAAACTAC  | 540  |
|    | CGGAATGAC CGCTATTGTA TGTTCATGTC TATTTCTCCT TTTCTAATAA CAATGTGTCT   | 600  |
|    | GTCAACCCAT AATGGTAACG TTTAATCATA TTAACCGGTT GATAATCLAT TTCTTCATGA  | 660  |
| 25 | TTGCTAAATT CACAACGAT GATACCATT TCTTTCAATA AATTAACCTC TGAATATTAGT   | 720  |
|    | TTTAAAGCTT TATCAATGAG ACCTTTTATA TAAGTGGAT CTAAAGAAAT GACATCAAAAT  | 780  |
|    | TGAATATCAC GTTTTGACAA TGCTTTTAAA GCTCTATCTG CATTATTTTT ATAAACTTCA  | 840  |
| 30 | GATGTGCGCT CTAAATCCAA ATTCCGAAGA TCTTGATTTAA TAACTTTTAC AGCTTTAAAA | 900  |
|    | TTTTGTGCAA CAAAGATTAT CTATTCCATA CTTGAGACGA GTGCTTCTAT TCCAAGCGCC  | 960  |
|    | CCGCTTCTCG CAAATAAATC TAAACCTATA CTTGACACAT CATATAAATC ATTAAGAGATA | 1020 |
| 35 | CCTTCTTTAA CTTTATCCAT AGTTGGTCTG GTATTACGCG CTTCACATA TTTCAAAGCT   | 1080 |
|    | TTACTTTTAT GTTTACCTGC AATGACGCGC ATGTGTGTCA CACTTCCAAT TCATTAGTT   | 1140 |
| 40 | ATTTAATATA ATTTATTGAG AAAAAGGAGA ATGATAAACC AATGAACAA ACATTATTATA  | 1200 |
|    | CACTTGTGGA AGGCTTAAAC GATTGTGTG AATTCTAGAC GATGATTGAA TATAACCATC   | 1260 |
|    | AACGTATTGA TAAATATTAT TATTTTCATT CACCACAAGC TGAATAATAA AAGTCATCTG  | 1320 |
| 45 | TAGCAATCAT TATGAACCTC ACAACTGGCA ATCATTTCOA AGCATTTTAT ATCATGATAA  | 1380 |
|    | ACGCTATTAA ATATCCATAC CCAGATTCAA ATAAAAAGTT TCAATAGATA AATGATTGTG  | 1440 |
|    | CTGAAAAATT CGACAATCCA ATTTTAGTATA TCGATGTACA GCGCCCTCAA GCATTTCAG  | 1500 |
| 50 | ATTATTGGTT ATATTATAAT TATTTATAAT GTGTGTTAAG GCTGCAAAAT TGGATACCG   | 1560 |
|    | aACTTCAATA ATATTAATTA TATATTTCTG GTTCTCTTT TCTGCAAGTT TCTTTAAGT    | 1620 |

TTTGATCAAC ATCTTGCTCA TTCACATACA TAATTACAAA TTTACGATCT CTATTTGAAT 1740  
GAACG 1745

5 (2) INFORMATION FOR SEQ ID NO: 503:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1035 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 503:

|    |             |             |             |             |            |              |      |
|----|-------------|-------------|-------------|-------------|------------|--------------|------|
|    | TCGCTCTTTAA | TCTTGCTGAC  | TTTAGAAGGC  | TTACAGAGTC  | GSTTACCATT | TTTTGCATCT   | 60   |
|    | TTAACTGATT  | GAACTAAAGC  | TtGACGTGTA  | GATTTATCAG  | CTAAACTAAT | TGCACCACCA   | 120  |
| 20 | ATTACGGCAC  | CAATTAAAAAT | ACCAGGAACA  | AAITTTATTTT | CCATAAAAAA | CTACCCTCT    | 180  |
|    | TTCAAATTTG  | CATCTTTTAC  | GATGTAGTCT  | ATTAAATTAT  | CACAAGATGA | TAATACCATG   | 240  |
|    | TCGTATACAC  | CTTCAAAATT  | ATTCTGTGTAG | TATGGATCTG  | GTACATCACT | CTCTTCCATA   | 300  |
| 25 | TTACTAAATT  | CTAACAGITT  | GAACAATTGT  | CCCTTAAGAT  | TAGGATTGAT | AGATTTAATA   | 360  |
|    | TTATCAACGT  | TACTTTGATC  | CATAGCCACA  | ATGTAATCAA  | AATCATCTGT | CGCTTCGAAT   | 420  |
|    | AATTCACATA  | TCATGCCATC  | AAATGGAATA  | TTGTGTTTGT  | TGAGAATTTT | TTGTGTACCT   | 480  |
| 30 | TCATGAGGTG  | GCTCTCCTAA  | ATTCCAGCTA  | CCAGTACCTC  | TTGAATGTAC | TTTAATATCA   | 540  |
|    | TGAATATTTC  | TGCTTTTAAG  | TCTTTGTGCG  | ATGATTGCTT  | CTGCCATTGG | AGAACGCAT    | 600  |
|    | ATATTGCCAA  | GACAGACAAA  | TGCTACATCT  | ACCAATTTGA  | TTCTCCAAA  | CTATGTAGTT   | 660  |
| 35 | ATATCCCAT   | TTTATAGCGA  | CTTTAAACAA  | TAAGAAAGCA  | GATTATATAA | AATTCATTATTA | 720  |
|    | AAGTTTATTA  | AATTTGTGATA | CTTTGATAAC  | ATAACTATTA  | TTAGAGGTGA | ACATTTGTGGC  | 780  |
|    | TATGACAAAT  | GAGAGAGAAAG | TnTTAGCTAT  | TAGAGAGAAG  | TTAAATAttG | TTAATCAAGG   | 840  |
| 40 | ATTATTAGAT  | CCCTGAAAAAT | ATAAAAATGC  | AAATGAGGAA  | GAATTAAACG | ATATATATGA   | 900  |
|    | TTTTTGKTCaA | yCAGAGAGAA  | GATTTGTCGCC | AAGTGAGTGT  | mCAGCTATTG | CTGaCGCTTT   | 960  |
| 45 | AGGACAATTG  | CGACACGAAT  | AGGAGTGGGA  | ATTTTGACTA  | ATTACAAAGA | AAAGTTACAA   | 1020 |
|    | CAATACGCTG  | AACTA       |             |             |            |              | 1035 |

(2) INFORMATION FOR SEQ ID NO: 504:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1284 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double



(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 504:

45 (2) INFORMATION FOR SEO ID NO: 505:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 5763 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

EP 0 786 519 A2

|    |  |      |
|----|--|------|
|    | ATAATTAACA TCTTTGCTA TATACCACCA GTTTGATACA TAAATATCG CAGCAATAAT    | 60   |
| 5  | ATCATGTTTA ACCCTAATGA TATTATCTGA TTTTAATAAT AAGGTTGCTG TCCCTACAAC  | 120  |
|    | CATTAAATAAA ACTATGACTG CTGGAATAAA ACGTTTAAAA CGACGTATCC AAAAGCTTTT | 180  |
|    | CAATTTAATG ATACCTGTGT CATCATACTC TTTGAGTAAT AAGCTTGTA TTAATAAACC   | 240  |
| 10 | AGAGATCACA AAAAATGTAT CCACACCTAA AAAGCCACCT GtCAACCATT GCITTATTAA  | 300  |
|    | GTGGTAAATA ATAATTCCTA GAACAGCGat TGCCCTCAAA CCATCGAGCC CTGGAATATA  | 360  |
|    | TCTCATTTTC TTATACCTTG TAAACCCCTT TGTTTGTTC ATTTTTCAT TCTTCCCTTT    | 420  |
| 15 | TAAAATGTT CTCTTAGAT GCTTAATTAA ATTTAGTTAT GCTGTTTAAA AGAATATTGA    | 480  |
|    | AATGCATATG TATATTATG AATTACGACA TCATCAAAAT CATATTGACT AAAATACTGT   | 540  |
|    | TAAATATAAA AAATTACCAA TGATGATTCT TACTTCGAAA TCCAATTGT AATGCAACTC   | 600  |
| 20 | CGCAACTTAA AAACCTGAA GTATTATGTA TTGTAATATA ACTGTAATAT AAATICAATT   | 660  |
|    | TATTATAAAA ATTTTCAAGA AAATATTCAA CTAGAAAATG AATTGTGCAC TCTTGAAGT   | 720  |
|    | GCAAGTCACT GTCTTAATTC ATATTTTGTG AAACAAGTTA GATATAAAAT TTCAAAAATA  | 780  |
| 25 | AATCAGAAAC TAGAACATAA ATAAGGCTCC CTTCAAAATT TTCATTTTTC AATGTCTACT  | 840  |
|    | TTGAAGGGAG CTATTTCACA ATGAATTATA CTCTACAATG TTATATTGAC TGGGGGCCCA  | 900  |
| 30 | AACACAGAGA ATTTGAAAAA GAAATTCCTAC AGGCAATGCA AGTTTATGTT AGCTCACACC | 960  |
|    | AAGTGAATC TTAGCGTAAC GTGACATCAT ATCTTTTGTG CAAGGTGGAC TCCATACGAT   | 1020 |
|    | ATTCACITCA GTATCTGAA TTTCAGGAAT CTCTGCTAAT ACTGTTTTTA CTGgaTCAAT   | 1080 |
| 35 | AATTTGAGGT CCCAATTGGAC ATCCCATTGA TGTTAAAGTC ATATCAACTG TACATAGGCC | 1140 |
|    | TTCAATCACTA ACATTCACTT TGTATACTAA ACCCAAATTA ACGATATCAA TTCTTAATTC | 1200 |
|    | AGGGTCAATT ACCATTCTTA ATGCACCTAA GATACTATCT TTCAATGCCT CTTCATCCA   | 1260 |
| 40 | TATCACCTCT TTAATGTCAT ATTATTCTATA ATATATCAAA TATCCGACAA AACGCCAATA | 1320 |
|    | AAATGCTATG ATGTATCTAT ATGAACTAAG CACTTATGA GGAGAGAGAT ATGCAACCAC   | 1380 |
|    | ATTTAATATG TCTAGACTTA GACGGAACAT TATTAAACGA TAACAAAGAA ATTTTCATCAT | 1440 |
| 45 | ATACTAAACA AGTATTAAAT GAATTACAAC AAGGTGGaCa CCAAATTATG ATTGCGACTG  | 1500 |
|    | GCAGACCTTA TCGTGCAAGT CAAATGTATT ATCATGAATT AAATTTAAAG ACACCAATTG  | 1560 |
| 50 | TTAATTTTAA TGGCGCTTAC GTACATCACC CTAAGAGTAA AAACCTCAAA ACTTGCCATG  | 1620 |
|    | AAATTTTAGA TTAGGCATC GCACAAAACA TTATTCAAGG ATTACAACAA TATCAAGTAT   | 1680 |
| 55 | CGAATATTAT AGCAGAAGTG AAAGATTATG TTTTCATTAA CAATCATGAT CCAAGATTAT  | 1740 |

EP 0 786 519 A2

|    |   |      |
|----|---|------|
|    | AAGAATCCCC TACCTCAATT TTAATTGAAG CCGAAGAAAG TAAATACCT GAAATCAAAA  | 1860 |
| 5  | ATATGCTTAC TCATTTTAT GCCGATCATA TTGAGCATCG ACGCTGGGGC GCACCATTC   | 1920 |
|    | CTGTCTATGA AATTGTAAAA CTGTGTATTA ATAAAGCAAG AGGCATTGAG CAAGTTAGAC | 1980 |
|    | AATTTTAAA TATTGACCGA AATAATATTA TTGCAITCGG TGATGAAGAT AATGATATTG  | 2040 |
| 10 | AAATGATIGA GTACGCGCGT CACGTTGTTG CTATGGAATA TGGTTTGCAG GAACCTAAAG | 2100 |
|    | ATGTAGCGAA CAATATTACA TTCAACAATA ATGAAGATGG CATTTGTCGA TATTTGAATG | 2160 |
|    | ATTTCTTTAA TTTAAATATT AGATATTACT GTTAATTTAT AACTAATCAT TTTATAATAT | 2220 |
| 15 | TTTAAACAA TAGGAGGTAA GTTACGATGC CCAAAATAGT CGTAGTCGGA GCAGTCGCTG  | 2280 |
|    | GCGGTGCAAC ATGTGCCAGC CAAATTCGAC GTTTAGATAA AGAAAGTGAC ATTATTATTT | 2340 |
|    | TTGAAAAGA TCGTGATATG AGCTTTGCTA ATTTGTGATT GCCTTATGTC ATTTGGCGAaG | 2400 |
| 20 | TTGTTGAAGA TAGAAGATAT GCCTTAGCGT ATACACCTGA AAAATTTTAT GATAGAAAGC | 2460 |
|    | AAATTACAGT AAAAAGTTAT CATGAAGTTA TTGCAATCAA TGATGAAGA CAAAGTGAT   | 2520 |
|    | CTGTATTAAA TAGAAGACA AACGAACAAT TTGAAGAATC TTACGATAAA CTCATTTTAA  | 2580 |
| 25 | GCCCTGTGTC AAGTGCAAAAT AGCCTTGGCT TTGAAGTGA TATTACATTT ACACCTAGAA | 2640 |
|    | ATTTAGAAGA CACTGATGCT ATCGATCAAT TCATCAAAGC AAATCAAGTT GATAAAGTAT | 2700 |
| 30 | TGGTTGTAGG TGCAGGTTAT GTTTCATTAG AAGTCTTTGA AAATCTTTAT GAACGTGGTT | 2760 |
|    | TACACCTCAC TTTAATTCAT CGATCTGATA AGATAAATAA ATTAATGGAT GCCGACATGA | 2820 |
|    | ATCAACCTAT ACTTGATGAA TTAGATAAGC GGGAGATTCC ATACCGTTTA AATGAGGAAA | 2880 |
| 35 | TTAATGCTAT CAATGGAAAT GAAATTACAT TTAAATCAGG AAAAGTTGAA CATTACGATA | 2940 |
|    | TGATTATTGA AGGTGTCGCT ACTCACCCCA ATTCAAAATT TATCGAAAGT TCAATATCA  | 3000 |
|    | AACTTGATCG AAAAGTTTC ATACCGGTAA ACGATAAATT TGAACAAAAT GTTCCAAACA  | 3060 |
| 40 | TTATGCAAT AGGCGATATT GCAACATCAC ATTATCGACA TGTCGATCTA CCGGCTAGTG  | 3120 |
|    | TTCTTTAGC TTGGGGCGCT CACCGTGCAG CAAGTATTGT TGCCGAACAA ATTGCTGGAA  | 3180 |
|    | ATGACACTAT TGAATTCAAA GGCTTCTTAG GCAACAATAT TGTGAAGTTC TTTGATTATA | 3240 |
| 45 | CATTTCGAG TGTCGGCGTT AAACCAAAAG AACTAAAGCA ATTTGACTAT AAAATGGTAG  | 3300 |
|    | AAGTCACTCA AGGTGCACAC GCGAATTATT ACCCAGGAAA TTCCCCTTTA CACTTAAGAG | 3360 |
| 50 | TATATTATGA CACTTCAAA CGTGAGATTT TAAGAGCAGC TGCAAGTAGA AAAGAAGGTG  | 3420 |
|    | CAGATAAAGC TATTGATGTA CTATGATGG CAATGATGAA CCAGCTAACT GTAGATGAGT  | 3480 |
| 55 | TAACTGAGTT TGAAGTGGCT TATGCCACC CATATAGCCA CCTTAAGAT TTAATCAATA   | 3540 |

|    |              |             |             |             |             |             |      |
|----|--------------|-------------|-------------|-------------|-------------|-------------|------|
|    | GTTCAGAAATTA | TGTTGGAGCTG | GTACTACTAT  | CCAGTCCATT  | TTTTATGTTT  | AACATTTTTTA | 3660 |
| 5  | GAATCAAAAA   | AGACATAAGG  | TCTTGGACTA  | ATAAATTGTC  | ATGCCTTATG  | TCAATATACTA | 3720 |
|    | TATGTCTTAT   | CAATTAGCCA  | ATACCGAATA  | ATTTTGATAT  | AGGSCCTAAC  | GGTAGAATGA  | 3780 |
|    | CACCTAATAC   | CATTGTGATG  | ATAAATTAATG | CAATTGTTAT  | CCAAAACAAT  | GTGTGACTTT  | 3840 |
| 10 | GTTCATGTCT   | CTTCTTTTTA  | GCAATCGACA  | CTTCCATCAA  | TCCAACCTACT | GCAACACCCAC | 3900 |
|    | ACAGACTTTT   | CAATGTAAAG  | AACATATATG  | TGCCCCGGAT  | ATTCATAAAT  | GACTGAATTA  | 3960 |
|    | ATATCCAAAA   | TCTGAAATT   | ACAGTCACAA  | GCAATAATAA  | CGGTAAAAAT  | ATGTGCAACG  | 4020 |
| 15 | GTTTGAAAAA   | TGGTGATCTG  | CCITGATTTT  | GCTAAATGTT  | TAGTATGTGA  | GCGATAAATA  | 4080 |
|    | AAATAATCGC   | TAATACCCAA  | CTTAATATAT  | GTAATGTAA   | CATACTGATT  | CCCCCCACTT  | 4140 |
|    | TAATTATTTA   | TATTATTAAA  | TTAAGCTTC   | TTGGGATTAA  | TACCCACTTG  | CTTGTAAATT  | 4200 |
| 20 | AATCATGATT   | TGATTATACA  | CGAATATATA  | TCTACCACA   | CTTCTATATT  | TGAGAGGAAG  | 4260 |
|    | AACATGACAT   | TTTATTCCCT  | ATTAGAATAT  | TGTGAATCTG  | CTGTAAAATA  | ATCAACTACT  | 4320 |
|    | TCTTAATATC   | AATATTTCAC  | TTTCATCTCA  | AAATGGTAAC  | ATTATAAATA  | ATTTATCTCT  | 4380 |
| 25 | AACACCTTTT   | TAGAAAAACA  | AGAAAAAACT  | AACCAATCTA  | TATAAAGACT  | GGTTAGCTTT  | 4440 |
|    | TTAAATGATA   | ATTATTTAGC  | GATATAAGTT  | GTCACGGTTC  | CAATATTATC  | AATAGTCACT  | 4500 |
|    | TTAACTTCAT   | CACCTGGATT  | TAAAAATTTA  | GGTGGTTGCA  | TACCTGCACC  | AACGCGCTGCT | 4560 |
| 30 | GGTGATACAG   | TGCAATAAAT  | ATCTCCCGGA  | TGATGTGCAA  | CATATTTTGA  | AATTTCTCTCT | 4620 |
|    | ATTAATTCAT   | CAATTTTAAG  | AATCATTTTC  | CGGTGTGTAC  | CTATTTGTCT  | AATTTCTAATG | 4680 |
| 35 | TTAACTTTTG   | TAACAATATT  | TACATTTTCA  | GGTAATGGTA  | GTTCGCTTTT  | AGTAACGATA  | 4740 |
|    | TAAGGACCCA   | TTGGGCAACC  | GCCAGTTAAA  | CTTTTGTGTA  | AAAATGCTTG  | ATCTTGTTC   | 4800 |
|    | CTTTGTGCTT   | TGCGATCAGT  | GATATCGTTA  | ATAAATTGTAT | AGCCGTAAAC  | ATAATCTAAA  | 4860 |
| 40 | GCTAATGCTT   | TTGGAATCTT  | TTCCACGAC   | TTACCAATAA  | CAATACCTAA  | TTCACTTCTCA | 4920 |
|    | TAATCTAATT   | GATCAGTAAT  | ATCTTTATGA  | TTTGGAAATG  | TTGCATTATC  | TCTGTGTTAA  | 4980 |
|    | GATGACGCTG   | CTTTTGTAAA  | TACATATAAT  | TTTTTCCACT  | CATGATTTAA  | TTCTGTTGCA  | 5040 |
| 45 | TGATCTTTGT   | AATTTCTACC  | AAAAGCAATC  | ACATTATTCTG | GAGGTGTTAC  | TGGTGGTAAA  | 5100 |
|    | AAATCAATGT   | CATTAAATGA  | AAATTTATAG  | TCTTCAGCTT  | TGCCCGCTACT | TTCTGTGCTGT | 5160 |
|    | ACAACTGCTT   | TAGTCACTTG  | TTCTTGAAAA  | CTTAAAGTAT  | GATTTTGTGT  | TAAACAGACT  | 5220 |
| 50 | AACAATGTTT   | TAGGATGGAA  | ATCTCCTCTT  | GCAAACTGAC  | CAAACTCTGT  | TGTTAAATCC  | 5280 |
|    | CATACAGCAT   | CTTCGCGTTT  | TACTTTAAAC  | CCATATGAAG  | TTTTGTCAAT  | ATACTTGAAT  | 5340 |

## EP 0 786 519 A2

TTCGTTATCA AATAACAAAT AAATAAGTAA GACAATTTTG AAAATGAGTT GTGTTCAATC 5460  
 TGCTACAAGG ACTTTGCACT TAATCGAAAT TATTTTTTAT TCTTTTGAAA ATCAAAATAC 5520  
 5 TATAGTTGCA ATGTACCAA TTGGAAGAAG TATAAATAAC CTTTAACTTC TTTATTAAAG 5580  
 ATCGTTTGAA GGGTATTTTG ATAATATTTC ATCTGTAICT TATATTATT TTTTAATTGT 5640  
 10 GTACCAATTT CTTCATCTGT CATCCACGG CGACGATTAA ATGCATCGGT TTTATAGTCT 5700  
 ACAAATAAT GCACACCATC TTTAACAAAG ATTAAGTCAA TCATACCTTG AATAATTGAG 5760  
 ACG 5763

15 (2) INFORMATION FOR SEQ ID NO: 506:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 422 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 20 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 506:

25 CCATATGATT TTGTGCAATA ACTCTTTTTC TTCTTCTTT TCGTAAAAAG TTGTACATCG 60  
 CTTTGTGTT TAAGAGACTA TTGTTTCTT TAGGTTTTTG AACTTCACTC AGTGATTTTC 120  
 TAGTTGTAA CACTAAAATT CCAACTGTTG TATCTTTGTA TCTAGCCATA ACTTTATTCA 180  
 30 GATGTTGTC ATTTGTAATA ACTACGACAT AATTAAACAC TTCATAATAA TCATTAATTT 240  
 GATTATCTAA TCTATCCAGC TTATCTAATT CTGTTTAAAT CTCATAGACA ACGCCTTTGC 300  
 35 TGTGTAACAA TATAAAATCA GCAATACTTT TCCCTATGGG CATCTCAGAA AGTGCAAGTAG 360  
 TTGTATTAAAT AGAATGTCGT CCTAGAAGGA GTkTATTAAG TATnGCTTTT TTGTAAAAAT 420  
 AT 422

40 (2) INFORMATION FOR SEQ ID NO: 507:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1188 base pairs  
 (B) TYPE: nucleic acid  
 45 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 507:

50 GCTTATGTAT TCTCAAATA TTTATGTGAT ACGCAAAGGA GACATGGTTA TTCGACCAGC 60  
 ATTTGATGAT GACGATCAAA GAAACGGTAG TGAAATAATT CGGTTTGACA AAACGCGTAT 120

55

|    |            |            |            |             |             |             |      |
|----|------------|------------|------------|-------------|-------------|-------------|------|
| 5  | TTATCTTGGC | AAGAAGCAG  | AGACAAACC  | CATTACTGGC  | ATTTCCTAGTA | AACCACCTAT  | 240  |
|    | TTTACTAACA | CCATTATTTC | CACTTATTTT | TTTCCCAACA  | CAITCTGACA  | GACAAATGA   | 300  |
|    | AAATATTTGG | TTAAATATGC | ATTATATCGA | AAGTATTAAG  | GAATTAAAA   | ATCGTAAATG  | 360  |
|    | TAAAGTGACA | TTTATTATAA | ATGAATCAAT | CAITCTTCAT  | GTTTCATACC  | ACAGTTTATG  | 420  |
| 10 | GCATCAATAT | AACAATTCCA | TTTTTTACTA | TTACATGGTA  | GATAAACAAT  | CTCGCATGAT  | 480  |
|    | ATCAAAAAAT | CCCACCAAC  | CAATAGATTA | TAATAAGCC   | ACATTGAATG  | TGTTTGAAGC  | 540  |
|    | ATTGACACGC | TATTCTTTAT | TTGAAGATAA | ATAAATGTG   | TATTTTAAAA  | ATATGGCGAA  | 600  |
|    | TGTTTTATAA | ATATAGTGTA | AATGTTCGCG | ATATTTTTTT  | AAGGTATCTA  | TTGCAAAATTA | 660  |
| 15 | ACTTAATCTT | GTTATAATAA | TATTTGTGCT | TGATATTTCA  | ACACATACAA  | ATTAATCCAC  | 720  |
|    | AGTAGCTCAG | TGGTAGAGCT | ATCGGCTGTT | AACCGATCGG  | TCGTAGGTTC  | GAGTCTTACC  | 780  |
|    | TGTGGAGCCA | TTGGAACGTT | ACTCAAGTTG | GCTGAAGAGG  | CGCCCTGCT   | AAGGTGTAGG  | 840  |
|    | TCGCGAAAGG | CGCGAGGGTT | CGAATCCCTC | CGTTTCCGTT  | ACTTGCTAAA  | ATGGTATATA  | 900  |
| 20 | CCATTTTATC | TTTTTTGTTT | ACTTATATAT | AATGAATGAG  | AAITTTCACTG | TTCTTTTATA  | 960  |
|    | TCAATTTTAA | AATTTCTAAA | ACCTTTCTTA | GATAATCTTC  | TCTAAGAAAG  | GTTTTTATAC  | 1020 |
|    | TTGTTGAACT | TATAATTAAT | TTATTACATA | GCAATTTACTA | CTGTGTTTTA  | ACTATAAAAT  | 1080 |
|    | TATCACTACA | TGAATATACG | TAAATCGGAT | CTCTTAACTT  | CTCTGCAATT  | AATGCACTCA  | 1140 |
| 30 | TGTGTTTCAT | CGTATGATTC | ATGTATAATA | GCATTTTCTA  | AATAATTC    |             | 1188 |

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 840 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

|    |            |             |             |            |             |             |     |
|----|------------|-------------|-------------|------------|-------------|-------------|-----|
| 45 | CCCAACTTCG | GTTATAAGAT  | CCCTCAAAGA  | TGATGAGGTT | AATAGGTTG   | AGGTGGAAGC  | 60  |
|    | ATGGTGACAT | GTGGAGCTGA  | CGAATACTAA  | TCCGATGAG  | ACTTAATCAA  | AATAAATGTT  | 120 |
|    | TTGCGAACAA | AATCACTTTT  | ACTTACTATC  | TAGTTTGTAA | TGTATAAAIT  | ACATTTCATAT | 180 |
| 50 | GTCTGGTGAC | TATAGCAAGG  | AGGTCACACC  | TGTTCCCATG | CCGACACACAG | AAGTTAAAGCT | 240 |
|    | CCTTAGCGTC | GATGGTAGTC  | GAACTTACGT  | TCCGCTCAG  | TAGAACCGTG  | CTTGACCATAA | 300 |
|    | TATTAATCCA | CAGTAGTGCTA | GTAGGTAGAGC | TATCCGGCTG | TACACGATCG  | CTCGTAGGTT  | 360 |

EP 0 786 519 A2

GGCGGTAACA CGGGTTCGAG TCCCGTAGgA GTCATACAAG CAGAAGTGAA ATATCGCTTC 480  
 TGTITTTTITA TTACATATIT ATTGTTGAGG AAGGTTGTCC GAGCTGGCCG AAGAGCACGC 540  
 5 CTGGAaAGTG TGTAGGCGTC ACAAGCGTCT CAAGGGITCG AATCCCTTAC CTTCGGTAAA 600  
 GGeGCTTAAA TTGGTTTtAC CCATTITTAAG CGCTATTTTT ATTTGGACT CAATCCCTTG 660  
 10 ATATATCTGC ATTTAGCTA TTATCCTCAT TTTTACACTT CTTATTtATT TATATCCATT 720  
 TAAATTTTT TAGCCACAAT GTGACTAATT TTTGATGAAT AATCCTAATT TTAGtCACAA 780  
 GATTTTGAAG TTTAGTCACA AAACAAATCA TTCAGATTTT TTTCyATAAA TTTAGTTTCA 840

15 (2) INFORMATION FOR SEQ ID NO: 509:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1373 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 20 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 509:

25 TTCTATGAAA CTATCAGCTG TTTCTTTTTG CTtGGTAAAA ATCGTTTTAA CGCCATGATA 60  
 CATGTATtCT CTTTCAaCAC CCGATTGGTC ATAAGCATCG AAAACACAAA TAATTTCAATC 120  
 30 TGAAATAACT GCATTATAAT TTGCAATTGC ATCTATTAAT TGCATTCTAG CTTCtCTTAA 180  
 ATTCTCTTTT GCAATGGCGC TTAGCGTTGG TGATTGtCCT ATCATATtGT ATCCATCAAT 240  
 GATTAAAGTAA CGTTCTTTCA TTATATTTCT CCAACATCAT GTCTTTTTCG AAATACTTCG 300  
 35 TACATCATTa AACTTGCTGC AACCGAAGCA TTCAAACtGT TTACATGTCC AACCAITGGA 360  
 ATCTTAATAT AAAAATCGCA TTTATCACTT ACTAGGCGAC TCATACCCTG TCCTTCGCTA 420  
 CCAATTACAA TAGCCAATGA CATGTCCGCT TCTAGATTTC TATAATCTGT TGCATTATTA 480  
 40 GCTTCAGTGC CAGCTACCCA AAAGCCATTa TCTTTTAGTT CATCGATAGT TTAGCTAAA 540  
 TTTGTCACTC GAATAACTGG TACATGTtCA ATTGCACCTG TTGAGGCTTT TGCAACTGTT 600  
 TGCCTTAGTG TAACtGAACG ACGTTTAGGA ATAATAACAC CATCAACTCC CGTTGCATCG 660  
 45 GCTGTCTTta AAATTGATCC CAAGTTATGT GGGTCTTCTa AGCCCTCTAA TATAAGTACT 720  
 GTCAATAAAC CTTCtTTTTT TTTTtGCTGT TTTAAAAATt GATCGAAGTC AGCATATTCA 780  
 TATGGTGCAa TAAGCGCTGC AACACCCtGA TGTGGTGcAT TTGCTAAAAA ATCTAAATTTA 840  
 50 GATTtTGGA CAGTTTGAAC AATGATTTTT TGATCTTTTG CATTTTTTAA AATTTCATTA 900  
 ATTTGTtGCT TTTTAATACC TTCTTGAATC AATATCTTAT TTAyCGGATG CCCAGTAATA 960

55

## EP 0 786 519 A2

TTCTGTTTACT ATTGTTATTA TTTTATTTAA TAATGCCTCT AATCGTTCTT CTCTTTTTTC 1080  
TAAATAaAGA AAACCTATCA CTGCTTCTAT CKCTGAACIT TTACGATATG TTtGAACATC 1140  
AGTGTTTTTA GCTTTAGTAT GACTTTTCGC GTTACGCCCT CGCTTCAAAA TATCCATTTC 1200  
TTCTGTCTGA AACCATTTCT GCTCCATTAA ATATTCTAAG GTTTGCGCCT GACTTTTGGC 1260  
AGATACATAT TTTTtAGACA TTTGATGTAG TTTATTAGGC TTACTTTTAA GCTTTAAAAC 1320  
GATATAGGTA CGTACATATT GATCTAaAGC TCGCTCnCCC ATATATGCTA AGG 1373

## (2) INFORMATION FOR SEQ ID NO: 510:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 717 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 510:

TCAAGTGCCA ACTACACAAC TGATTTACAT TCTTTAGGTC AATATGTACA AGAAGGCCGT 60  
CGCTTCTTAT TCGAAACAGT TGTAAAAGTA AATCATCCTA AATATGATAT TACTATTGAA 120  
AAAGATAGTG ATGATCTAGA CGGATTAAAT TATTTAGCTG GTAAAACAAT CGACGAaGTT 180  
AACACAAAAG CATTCGAAGG TACATTATTA GCGCATACTG ATGGTGGTGT TCCTAACATG 240  
GTAGTGAACA TTCCACAATT AGATGAAGAA ACTTTCGGTT ACGTCGTATA CTTCCTCGAA 300  
CTTGCTTGTG CAATGAGTGG ATACCAATTA GCGGTAAATC CATTTAACCA ACCTGGTGTA 360  
GAAGCATATA AACAAAACAT GTTCGCATTA tTAGGTAAAC CTGGTTTTGA AGaCTTGAAA 420  
AAAGtATTAG rAGAACGTTT ATAAAATACA TTACTTCAAA GATTAGTGAA GTTTGAAAAG 480  
ATAGAAGTAG ACGTTAACTA TTTAAAGCAT ATTTTCGAGG TTGTCATTAC AAATGTAAAA 540  
ATGTAATGAC AACCTCGITT TTATTTATAT GCAAGAATA GGTACTAGC TAaTGTGACA 600  
AGATGTTAAG AGAAAATTAA aGATAAAAATA ATATCTGCCT TACAATAATA TTGTTATACT 660  
ACTAGnGACT GATTTATTAG CATGATTACA TGTTAATGTT TCtTTACTTA GTAATTA 717

## (2) INFORMATION FOR SEQ ID NO: 511:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 2700 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear



EP 0 786 519 A2

|    |  |      |
|----|--|------|
|    | AATCTAATTT TTCCTCTAA TCTCGGATTG GATAGTGGTC AATTTTCTG GCCCGCCATT    | 60   |
| 5  | TTAGCGTTTG TTCTAACTGG GATTGGTTTA CCATTATTAG GTGTGATTGT AGGTGCATT   | 120  |
|    | GATAAAGAAG GATATATTGG CGCATTAAAT AAAATTTCAC CTAATTTTC ANATTGTTTC   | 180  |
|    | TTAATCATCA TTATTTGAC TATAGGACCA CTTTITGCAA TACCTAGAAC TGCATCTACA   | 240  |
| 10 | TCTTTTGAAA TGACAATTAC ACCAATTATA CATAGCAATA GTAGTATCGC TTTATTTATA  | 300  |
|    | TTTACGATT TCTACTTCAT AGTCGTTTG TATATTGTT TAAATCCATC TAAGTTAATC     | 360  |
|    | GATCGTATTG GTTCATTATT AACACCATT TATTGATTA CTATTTTAGC GATGATTATT    | 420  |
| 15 | AAAGGACTACT TAGACTTTAG CGTAATAGT GCTGGAAAGG GCAATGAAGC ACTATATCAT  | 480  |
|    | TCTAATTTTT CAAGTTTTGC TGAAGGCTTT ACACAAGGCT ATTTAACAAT GGATGCCATT  | 540  |
|    | GCAGCAATTG CTTTTCAAT GATTGTTGTT AATGCAGTAA AACTAACAGG CATTAATAAA   | 600  |
| 20 | ACAAATCAAA TATTCAAACA AACTTTGACT GCTGGTTTAA TTGCAGCCGT AGCTTTAATT  | 660  |
|    | TTTCATATATA TTTCAATTAG TTATATTGGT AATCATATGC CAGTAAGTGA CATGACGTTA | 720  |
|    | GATCAATTGA AATCCAAGA TCGAAACATT GGGACATATT TATTAACGAC AATGGCTTCA   | 780  |
| 25 | ACAGGATTTG GTTCATTGCG AAAATATTTA TTGGGCATCA TTGTGGCGCT GGCATGTCTA  | 840  |
|    | ACTACAGCAT GCGGGCTTAT TGTGTCAGTT TCTGAATATT TCCATAGAAT CGTACCTAAA  | 900  |
|    | GTATCATACA AAGCATTTGT ATTAGTTTTC ATTTAATGA GTTTTATTAT TGCTAACCAA   | 960  |
| 30 | GGTTTAAATG CTGTTATCTC AATGTCAATT CCGGTATTAA GCATTGTATA CCCAGTAGCA  | 1020 |
|    | ATAACTGTTG TATTATTAAAT TTTAATTGCC AAATTCATAC CGACAAAACG CATTTCACAA | 1080 |
| 35 | CAAAATCCAG TTATTATCGT ATTTATATTG TCGATTTTCA GTGTATTAGT TAAGTTAGGT  | 1140 |
|    | TGGCTGAAAA TTAACTTTAT AGAATCATTG CCTCTAAGAG CGTATTCCTT AGAGTSGKTC  | 1200 |
|    | CCAGTAGCAA TTATTGCAAC GATATTAGGC TATCTAGTGC GCATATTGTT AAAACAAGAT  | 1260 |
| 40 | CCAATTAAAT ATCAACAGGA ATAACGAATA ATATAAAGA GGTGGGACA TAAATCCCTA    | 1320 |
|    | AAAAACAGC AGTAAGATAA TTTTCAATTA GAAAAATCT TACTGCTGTT CTCTATTTAT    | 1380 |
|    | ACAATACTTC GTATTGAATG GCTTCGCTTT CCTAGGGTGC CGTCTCAGCC TCGGTCTTCG  | 1440 |
| 45 | ACTGGCACTG CTCCTCAGG AGTCTCGCCA TTAATACTAC GTATTAACTG GTAATTTTAC   | 1500 |
|    | TTTGAATATC TTTAAAAA TAAGACACTT TGCCCACTT GCACATAAAT GTAATAATCA     | 1560 |
|    | ATAAAATAAA TTTTCTGTGT TGGATCCCTT CGTATAATTT AATAAATACT ACTAACTAA   | 1620 |
| 50 | ATTAAACAGG TGCCTTATGT ATAAAAATTA TAACATGACC CAACTTACAC TACCAATAGA  | 1680 |
|    | AACCTCTGTT AGAATCTCT AAAATGATAA TACGCGATAT GTTAATGAAA TTGTGTAAC    | 1740 |
| 55 |  |      |

AAGATATGCA TACCGTAATG ATAGATATAG TTTTAAACGT GACTTCAAGC TATATGAATG 1860  
 5 TGATGACTGT TCATCATGTT CTTTGAGACA TCAATGCATG AAGCCAAATT CGAAATCCAA 1920  
 TAAGAAAAAT ATGAAGAATT ATAATTGGGA ATACTTTAAA GyCCAAATTA ATCAAAAGCT 1980  
 TTCTGAACCA GAAACGAAAA AAATCTATAG TCAAAGAAAA ATTGATGTAG AGCGTGTGTT 2040  
 10 TGGATTATG AAGGCTATTT TGGGTTTCAC TCGAATGTCA GTTCGAGGAA TAAATAAAGT 2100  
 TAAACGAGAG CTAGGTTTTG TATTAATGGC ACTTAATATA AGGAAAAATG CAGCTCAACG 2160  
 AGCTGTACAT TATAAAATAC ATATCAAAAA AGCTGATTC TATCAATAAA TTAATAGAAA 2220  
 15 TCAGCTTTTT TACATTGCCT AAGAATTAA TGTCCTAAGC CCTTTTATC GAATAACTTA 2280  
 TTGTAACCT TGTCTTTCTT GGTATTGTT TCGTTATTT TTTTCGTGT TTTGTTTCCA 2340  
 CTCTTTTTGA GTCAATTACAT GGTCAACTTG CATGTTAACT TCAACTACAT CTAAACGAGT 2400  
 20 AATATATTTA ACTTGTCTT TAACTAATTC AGTTACTTTA CGGAAGATT TTTGGTCTGA 2460  
 TTCACCATAT TCTAAGATTA CTTTTAAGTC TACAGCAGCT TGTTCCTAC CAACTCAAC 2520  
 AGATACACCT TGAGTAACAT TGTGGCCACT TGAGAATGCA TTAGTGAATG TATCAGTTAA 2580  
 25 GCCACCTTTC ATATCTAAGA TACCTTTAAC TTCACGTGCA GCGATACCAG CAATTTTTTC 2640  
 AACAACTTCA TCAGAGAATG TTAATTGnT TTTGAATTGA GGTCTCTGAT TnGTTCTnTG 2700

## (2) INFORMATION FOR SEQ ID NO: 512:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1135 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 512:

ATCAAAATACA ATAATTAAAA TAaTGGCTAA TACaCCTAAa ATaATAATcG AGTTaGAGTA 60  
 GCCTAAGCGA yCACcCTAAC aGTaFACATT CTAGGCATAT ATTiaTCTIT AGACATTGAA 120  
 GCGCTAACA TCGGAAAAAG TGTAATCCA GTGTTGGCTG CTAATACCAA AATCATCAG 180  
 45 GTCGTTGCCT GTACAAAATA GAAGGCAGCA TTATACCAA ATATTGTCAT AGCTAATTGT 240  
 GATAAACTG TCGTTTCGCT TTGTGGcAAA ATTCCAATAA CATATGCTAA ACCAAGGATA 300  
 CCACTAATA AAAAGCTAA AATTGAACCC ATAGCAATTA ACGTTITCAG AGCATTTTTTA 360  
 50 GCACTTGGCT CTCTAAaATT GGTACCAGCA TTGAAATAG CTTCAACACC CGTTAACGAT 420  
 GAAGCCCCTG ATGAAAAATG TTTTAATaC AAGAATAATG TTACTCCAGG AACCCGAGTT 480

|    |             |            |             |            |             |            |      |
|----|-------------|------------|-------------|------------|-------------|------------|------|
| 5  | AATATTAATA  | TCCTAACCC  | TATAATGAA   | AGATATACTG | GATAGGATAA  | TACGGTGGCA | 600  |
|    | GATTCAGTTA  | AACCACGTAA | ATTTAATATT  | AAAAATAAAA | GTACAGTAW   | ACATGCAATC | 660  |
|    | AGTCTCKTTAT | GCCCATATAA | ACTTGGGAAT  | CGACCAACAA | ATGCATCAGC  | ACCAGATGAT | 720  |
| 10 | ATCTCAACAG  | CGACAGTCAG | TATGTAAATGT | ACTAATAATG | AGCCCTCTGC  | AAGCAATCCC | 780  |
|    | CATTTTCTC   | CTAAATTGGT | CTTGACACCC  | ATATACGCGC | CGCCACCTTT  | AGGATATGCA | 840  |
|    | TAAATAATTT  | GCCTATAAGA | CATAATTAAA  | CGACGTAAAT | AGATCAGAAG  | AGCACCTGCA | 900  |
| 15 | ATCGGTAAAG  | TATACCAAGT | TGCAACTGCA  | CCCACTACTG | ATAGTGTAAAT | CAGTATTTGT | 960  |
|    | TCGGGACCGT  | AAGCCTACGA | AGATAATGCA  | TCGGACAGAA | GAATCGCTAA  | CCCTTTAAAC | 1020 |
|    | TTGGATAATCT | TTTCGCTCTT | TAGTTCCTG   | TTTTTTTATG | TGTCGCCATAT | ATAAAGCTCT | 1080 |
|    | TTAAATTGAT  | KGACATAATC | TACTATTCTT  | TTTTATTATG | TGATGAGAA   | ATCAC      | 1135 |

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 918 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

|    |            |             |            |            |            |            |     |
|----|------------|-------------|------------|------------|------------|------------|-----|
|    | TTATGGTAAA | ACITTAAACAC | AAGCATTACA | TAAATCGCAA | TCATTTATGT | GGCAAAAACG | 60  |
|    | TGGAATTTAA | GCCATTCTTA  | TTCAITTTAC | GTAATTATTA | GCCGTATATG | TGTAATAATA | 120 |
| 35 | CACATTTTAT | TCAGATTTTT  | TATCGGCGTC | CATTAAATCT | TTTAGCGATT | CTTTTACTGA | 180 |
|    | GATATTTTCA | ATAATAACTC  | TATATAATCA | ATTTGTAATT | GGCATATCCA | CATTTTITTC | 240 |
|    | TTTAGCTAAA | TGATAAACTG  | ATTTAGTTGT | ATTAATACCT | TCCAAACCA  | TATTCATTTT | 300 |
| 40 | AGATAATGCT | TGATCCAATT  | GATCACTCTG | TCCAAGTTTA | TATCTCAATG | TGAAATCCG  | 360 |
|    | AGAATGTGT  | GATGTGCAAG  | TAACGATTAA | GTCACCGATA | CCACCTAAAC | CTAGAAATGT | 420 |
|    | CATAGGATCG | GCACCTAACT  | TTTCACTTAA | TCTACTAATT | TCCGCTAAGC | CACGAGTCAT | 480 |
| 45 | TAATGACGCT | TTTGCAATTAT | CACCGTAGCC | AATTCAGACT | ACGATACCAC | TTGCTACTGC | 540 |
|    | GATGATATTT | TTCAATGCAC  | CACCAAGTTC | AACACCAATC | AAGTCATCAT | TCGTGTACAC | 600 |
|    | ACGCAAAATG | TCATTTCAATA | ATAAATCTTG | CGTTAATTTT | CTTACACTTT | TATCTTTTGA | 660 |
| 50 | TGAAGCAGCA | ACTGTAGTTG  | GTGCTTGACG | TAAACACTTC | TCGCGCATGC | TTGCGCTTGA | 720 |
|    | CACACGCGCA | ATACCTGCTG  | TATATTCAGG | TGAATAAGAA | TTCGTAATCA | TTTCTGCAC  | 780 |

CAGCTTATCA TTAATTTGAG AAGCAACTTC TCGCATTGCT TTAGTAGGTA AAGCCATTAA 900  
 GTAAATATCT GCAAAATTG 918  
 5 (2) INFORMATION FOR SEQ ID NO: 514:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 587 base pairs  
 (B) TYPE: nucleic acid  
 10 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 514:  
 15 CAATACTCGG TATTATAGAT AGTCCTACTA CATTTATATG GGTITGTATC ATTACATTAA 60  
 TTGCACATCA ACTTGAAGGT AATATCATT A CACCAATGT AATGGGTAAA TCTTTAAGTA 120  
 20 TCCATCCTTT AACAAATTATC GTTGITATTT TAGCAGCAGG AGATTTAGST GGCCTTACAT 180  
 TAGTTCGTGAT TGCAGTGCCA TTATATGCTG TACTTAAAC GGTGTITAGT AATATTTTCA 240  
 AATACGCCCA ACGCATTATT GATAAAGCAA ACAGTAATGT TAAGACTAA TTCTGTGGAT 300  
 25 GTCTTTTAAG AATATAAGAT ACTATCGCAT CAAAAGTTGA AACTACAGCT TTTGAGGCGA 360  
 TTTTTTTGTG CATAAAAAAT CAGTCmAAAT AATATACAAA TAATTTTCCA TCAGTCGGAT 420  
 TATTATAAAA GCAAAAAAGC TTTGCTCACA TATATAATAA CGTGAACAAA GCTGTTGAAT 480  
 30 GATATTATTT AATTGCGTGS AATCCGCTAT CTACATGAAT ATTTTCACCT GTAACGCCAC 540  
 TTGATAAATC ACTTAATAAG TAAGCCGCGAG TTTTACCTAC TTCTACT 587  
 35 (2) INFORMATION FOR SEQ ID NO: 515:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 812 base pairs  
 (B) TYPE: nucleic acid  
 40 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 515:  
 45 GGTTCATTCC AAAAAAGTAC GCGATAATTA GCGAAGAAGA ATTTGAAAAA TTAATGTTG 60  
 TTAAACCAAA TAAAAATAAT GTTTTCTGGT CAGTTATAGG AAGTTCGGTT TTGTTGGAG 120  
 50 TTACTTTAAG GAAATACATA CATGTTTTTG ATGTTCAATT AGATAAGCTA GTTGTAATGA 180  
 TATTGTGTGC TCTCGCTTTA ATTTGTGTTA TAGTTTTTTA TTTTAACTTA AATAGAAAGC 240  
 TTAAGTTAAA AGTGTTTGTAT ACAAATATTG AAAAAATATA GAGAGTTATA TAAAwacCAA 300

## EP 0 786 519 A2

TTTCATTAAT TGCCTTATG ACAATCGAAC CTCAAAATAT AATAATATTT ATTTATIGGA 420  
 TTATGATGAC AATGCTTTTC TTTTGTGTA AATAGACTTC GATAGGTAAAT GAAAAAGTTC 480  
 5 GCGTTATAAT GAAAAATAAT TGATTACATT TAAAAATATC TAAATGTGT CGACACAATC 540  
 CTTTAAAGAC GCTAGTAGAA TTTAAATGAC TTCTAATGTA TATGAAAGTG TATCAATATA 600  
 10 AAACCAATTG AAAAGAAGTG GAGACATTGC TTTGTGAAAC TGAAAAATAT AATAAGAATC 660  
 CCAATATAG AATTATCAAA TACAAAGATG AATATTTGAT GATTGATTTA GTAAGTACAT 720  
 GGTTAGCACT CTTTTTCCCA ATGATTAAAT GGTGATTCC AAAAAAGTAC GTCAAAATCA 780  
 15 GCGAAAAAGA TTTTGAAACT TTAACATTG TG 812

(2) INFORMATION FOR SEQ ID NO: 516:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 526 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 516:

GTTCTATTAT TTCTTCCAAT GGTAAATGGAT TTTCGTAAAA TGATTGATAT AAGTTGATAA 60  
 30 TCTCAATAAC TTTTAATIGA TCTGGAAACA TCGTTTTTGT GAACATTATA CTGCGATTTT 120  
 CACTTTGTAA TAACCTGTCT TTATCAAATA TCTCACCAGA ATTAGCATTA ACATTACCAA 180  
 TTAATATATC AATTAAAGTT GACTTTCAG CACCATTTTT TCCAATTAAA GCGATACATT 240  
 35 TACCTTGTTC AATATCGAAC GAAATATTTT TTAGAACACA TCTTTTATTA AATGACTTGT 300  
 TGATATTAGA TATTGGAATC AATTAAATCA CCTCTATTTT TTCTTAATT TAATATTAGT 360  
 AAATTTATTA GATTTAAAT AGAACACTTT GTCATAGATT TGAAATGACA AATGTCAATTA 420  
 40 TTAGATKAC ATAATATATT TATCGTATT TTAATTTTGG GCAAAATAAA AAGAGCCTCT 480  
 ATAATCGc TCCTTACAAA TAAATTATTA AATGGCGAA cTAAAT 526

(2) INFORMATION FOR SEQ ID NO: 517:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 454 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 517:

## 1443

|    |             |             |            |             |            |             |      |
|----|-------------|-------------|------------|-------------|------------|-------------|------|
|    | GGAAACACCTA | AAGAAAGAAA  | TGATGCATTA | AACACAGAGG  | CTGATATCTA | TGTAACCAAT  | 120  |
| 5  | AAAGAAAATA  | CTAAATGGTT  | ATGCGATCAA | TATAAAAAAG  | AATGGCCATT | TGATATGGTT  | 180  |
|    | GTAATTGATG  | AACGTGCTAC  | ATTAAAAAGT | CCTAAGAGTC  | AAAGGTTTAA | ATCTATTAAA  | 240  |
|    | AAGAAATTAC  | CACCTCATTA  | TAGATTTATA | GGATTAAACAG | GAACACCTAG | CCCAAAATAGT | 300  |
| 10 | TTACAGGATT  | TATGGGCTCA  | AGTTTATTTG | ATAGACAGAG  | GCGAAAGACT | TGAGTCTTCA  | 360  |
|    | TTCACTCGTT  | ATCGAGAAAG  | GTACTTTAAA | CCACACATC   | TATGTAGCGA | ACATGTTTTT  | 420  |
|    | AACGTGGGAG  | TAGAGAGCGG  | ATCTAGAGAA | AAGATATATG  | AACGAATAGA | AGATATATGT  | 480  |
| 15 | TTAAGCATGA  | AAGCGAAAGA  | TTATCTAGAT | ATGCCTGACA  | GAGTGTGATC | TAAACAAACA  | 540  |
|    | GTAGTCTTAT  | CTGAAAAAGA  | AAGAAAAAGT | TATGCAGAAT  | TAGAAAAAAA | CTATATTTTA  | 600  |
|    | GAATCGGAAG  | AAGAAGGAAC  | AGTTGTAGCT | CAGAATGGGG  | CATCATTAAG | TCAAAAACCTA | 660  |
| 20 | CTTCAACTAT  | CTAACGGTGC  | AGTTTATACA | GATGATGAAG  | ATGTAGAGCT | TATACATGAT  | 720  |
|    | AAGAAGTTAG  | ATAAGTTAGA  | GGAAATTATA | GAGGAGCTCT  | AAGGCCAAC  | AATATTATTG  | 780  |
|    | TTTTATAACT  | TCAAAACATGA | TAAAGAAAGA | ATACTTCAAA  | GGTTTAAAGG | AGCAACACCA  | 840  |
| 25 | TTAGAGGATT  | CAAACATAAA  | AGAAGCTTGG | AATAGTGGAG  | ACATTAAAGT | GCTTATAGCA  | 900  |
|    | CATCCAGCAA  | GTGCAGGGCA  | TGGATTAAAC | TTACAACAAG  | GTGGGCACAT | TATTGTTTGG  | 960  |
| 30 | TTTGGACTTA  | CATGGTCATT  | GGAAATTATC | CACACAGCAA  | ATGCAAGATT | ATATAGACAA  | 1020 |
|    | GGACAAATTC  | ATACGACTAT  | TATTCATCAT | ATTATTGACG  | ATAACCAAT  | AGATCAAGA   | 1080 |
|    | GTATATAAAG  | CTTTACAAAA  | TAAAGAACTA | ACCGAAGAAG  | TAAATGATGA | AGCTATTAAA  | 1140 |
| 35 | GCAAGAAATG  | CTAAGCATAA  | GTAATTGGAG | TATAAGATGG  | GAAAGGCATC | ATACGATATT  | 1200 |
|    | AAGCCAGGTA  | CATTTAAATA  | TATTGAGTCA | GAGATATATA  | ACCTACAAGA | GAACAAGAAA  | 1260 |
|    | GAGATAAATA  | GATTGAGAAT  | GGAGATACTT | AACCCACGA   | AAGAGCTAGA | CACTAACATT  | 1320 |
| 40 | GTGTATGGAC  | CGTTTGCAAA  | AGGTGAACCA | GTTGAACAA   | CTGAACTAAT | GGCAACAAGG  | 1380 |
|    | TTATTGACTA  | ATAAGATGTT  | ACGAAACCTA | GAAGAAATGG  | TGCAAGCAGT | TGAAAGTGAA  | 1440 |
|    | TACTTAAAGT  | TACCTGAAGA  | TCATAAGAAA | GTAATTAGGC  | TAAAGTATTG | GAATAGAGAT  | 1500 |
| 45 | AAGAAGCTAA  | AGATAGAGCA  | AATAGGAGAT | GCAATGTCAC  | TGCATCGTAA | TACAGTTACT  | 1560 |
|    | ACTATAAGAA  | AGAACTTTGT  | TAAACCGGTA | CGWATCATCG  | CAGGTATCAA | ATAACATTGT  | 1620 |
|    | CGAAGAGATTG | TGCAAAAGGC  | CTACAAATCT | GATCTAATAT  | TATAGTATCG | GAAGAAGTGA  | 1680 |
| 50 | TAAAGTTATC  | TAAAGTTTAT  | ACGACACAAG | TAGACAGGAC  | ACTACGCTAT | CGGTGTGTCT  | 1740 |
|    | TTTGTATATG  | AATCAAAGAG  | GTGTAAAGAA | TGACCAAGCA  | TAAATAACAT | TATAAGCATG  | 1800 |

55

EP 0 786 519 A2

AACACATTCT TGGATTCTTA AGCATAAAGT TGAATATTCT AACGAAAAGA TACCCTATAT 3720  
 AGAATGGGAA GAAGACGGAT TACTAACAAAT ACAAGATAAT CCTTATATAG ACTACCAAGA 3780  
 5 TGTITTAAT TGGATAATAA AGATGAATGA GCATTATGTT GTCGAAAAAA TCACCTATGA 3840  
 TAGGGCGAAT GCTITTTAAAT TAAATCAAGA GTTAAAGAAT TATGGCTTTG AAACAGAAAG 3900  
 AAcwAGACAA GGGGCTTTGA CCTTGAGCCC TGCaTTGAAG GATCTAAAAA AAATGTTTTT 3960  
 10 AGATGGGAAA ATAATATTTA ATAATAATCC TTTAATGAAA TGGTATATCA ATAATGTTCA 4020  
 GCTGAAACTA GACAGAAATG GGAAGTGGCT GCCATCTAAA CAAAGCAGAT ATCGTAAAT 4080  
 15 AGATGGTTTT GCAGCATTTT TAAACACATA TACAGATATT ATGAATAAAG TTGTTTCTGA 4140  
 CAAGGOTGAA GAAACATAG mATTATTAG TATTATAGAT ATAATGCGTT AAGGAGGTGA 4200  
 ATGTTATCGC AAAAGAGAAAT ATTGTCACAC GCATAAAGAA AAAATTGATA GACAATTGGA 4260  
 20 TTGaTCAGTC AGCTTCTAAG CTTTATGACT TTAGCCCATG GAAAAATAAA TCTTTTGGG 4320  
 GTGTAATCAA TAATACGCTT GAAACTAATG AAACGATATT TTCAGCTATT ACnAAGTTAT 4380  
 CTAATTGATG GCCTAGTTTG CCCTTGAAAA TGTATGAAGA TTATAAAGTA GTTAATACAG 4440  
 25 AAGTATCTGA TThACTTACA GTGTCACCGA ATAATTCTCT GAGCAGTTTT GATTTTATTA 4500  
 ATCAAATTGA AACAACTAGA AATGAAAAAG GTAATGCATA TGTG 4544

(2) INFORMATION FOR SEQ ID NO: 518:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 881 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 518:

CATGGCAGAA TATTGAAGCT GTAAAAAAG GACATGTAAT TTCATATAAA GCAGAAGATT 60  
 ATTGGTTCAC AGATCCTATT ACATTAGAAC ATTTGAGAAG TAAATTAAAA AAAGAAATTT 120  
 TAAATAAAA ATAATAGAAA TAAGTTGTAA AAATTTTCTT ATGCAITGGT ACTAATGTTT 180  
 45 TTAAGGAGTG ATTAATGAA GCAACTGGTT GGAATCCCG AATCAATGTT AATTCCTTTG 240  
 ATAGCTCGAG CAAAAGAGTA CGAAAACGAA AAACCAATAA TAAAAGACGC ACTATCARAA 300  
 AAAATATTGT ATGGTTTAGA TGATATGTAC AAAAATGTTA CATGTGATGA CATGTCCTCA 360  
 50 ATTGGAATTA GTATACGTAC TGTGATAATA GATTGTGTTA CTAAGAGGCT TATCAAGGAT 420  
 AATAAGATT TAATCGTGGT CAATATAGGT TGTGGCTTAG ATACAAGGTT TCAAAGATT 480



ACATTITTTT AAGAAAGTAA TAGTTATAAG ATGATATCTA AATCTATGCT AGATTACAGT 600  
 TGGATTGATG ATGTCAAAAA TTATAAATTT TTTAATAGTA AGTCAGATAT ATTGTTTATT 660  
 5 ATTGAAGGTG TATTGATGTA TTTTGATGAG AGTGAATGA CTCAAITATT GGACACTATT 720  
 ATCAAAAAGA TGGGAGATCA TAATTGACA TTTGCGATTg AATTTTGCTC aaaaaCAATT 780  
 10 GCGAATAATA CmaAGAGACA TCAATCGGTA TCCAAGTTAT CCTCACCACC TGTTTTTTAA 840  
 TATGGGTACC ATGATTITAn AAAATTGGAT GAnATTTACC C 881

(2) INFORMATION FOR SEQ ID NO: 519:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 3122 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 519:

TGAATaAAAA TATATTAATA GATAAACACA AATGTGTCCa AATACCCTTA GAGGTATTTG 60  
 25 AChAGTTCCA TCCAACIGTT TAAAAATACC CTACAGGTAT TTTAGGGAG GTTATTATGA 120  
 AACAAATCGG AGAAAaGTTT ATCGATGAAT TTAGTAAAGC AGAATTGGAA AAACATAGCCA 180  
 30 AGCAAGGGCA ATTAATTGAC GTTAGAACAG AAGAGGAGTA TGCATTAGGA CATATCAATG 240  
 GTTCACACT TCATCCTGTT GATGAGATTG AGTCATTCAA TAAAGAAAAA AATAAAACCT 300  
 ATTATGTAAT CTGTAGAAGT GGTAAACAGAA GTGCTAATGC TAGTAAATAT TTAGCTAAAC 360  
 35 AAGGTTATAA CGTTATAAAT CTGTATGGTG GTTATAAAGC TTATGAAGAA GAAAACGATA 420  
 GTTATGATAC ACAAGAAGAA TATAAAAGTA TAGAAATTAA AGCAGATCGT AAACAATTTA 480  
 ACTATCGTGG TCTTCAATGT CCAGGGCCAA TTGTAAAAAT TAGTCAAGAA ATGAAGAATA 540  
 40 TTGAAGTAGG TGACCAAAAT GAAGTCAAAG TCACAGACCC TGGATTCCCT AGTGACATTA 600  
 AAAGTTGGGT GAAACAAACA AGGCATACCT TAGTTAAGCT TGATGAAAAA AACAATGGAA 660  
 TTAATGCGAT TATTCAAAAA GAAAAAGCAA AAGATTAGA TATAAATTAT TCTGCTAAAG 720  
 45 GTACTACAAT TGTATTATTT AGTGGAAGAT TAGACAAAGC TGTAGCAGCG TTGATTATTG 780  
 CAAATGCTGC TAGAGCTGCT GAAAAAGATG TAACTACTTT CTTACTTTTT TGGGGGCTTA 840  
 50 ATGCATTAAA AAAAGTGCAA ACAGTTAATG TTA AAAAGCA AGGTATTGCA AAAATGTTTG 900  
 ATTTAATGTT GCCCAAAAAG AATATACGAA TGCCCTCTTC CAAATGAAT ATGTTTGGTT 960  
 TAGGAAATAT GATGATGCGC TACGTAATGA AAAAGAAAAA TGTGTATTCA TTACCAACAC 1020

EP 0 786 519 A2

|    |            |            |             |            |             |             |      |
|----|------------|------------|-------------|------------|-------------|-------------|------|
|    | TCATGGGTAT | TCAGAAAGAA | GAACCTAGAG  | ATGAAGTTGA | GTACGGTGGT  | GTAGGCACCT  | 1140 |
| 5  | ATATTGGTGC | TACTGAAAA  | GCGAATCATA  | ATTTATTTAT | CTAATTAATA  | CTATTATATA  | 1200 |
|    | AAGGAGTTGT | TATCATGTTT | TTTAAACAGT  | TTTACGATAA | TCATTATATCT | CAAGCATCAT  | 1260 |
|    | ATTAGTGGG  | TTGTCAACGT | ACAGGAGAGG  | CAATAATAAT | AGACCCCTGTT | CGTGATTAT   | 1320 |
| 10 | CGAAATATAT | AGAAGTTGCA | GATTCTGAAG  | GTTTAACAAT | TACACAAGCT  | ACAGAAACAC  | 1380 |
|    | ATATTCTATC | TGATTTTGCT | TCAGGAATTC  | GTGATGTGGC | TAAACGCTTA  | AATGCAATA   | 1440 |
|    | TATATGTGTC | TGGCGAAGGT | GAAGATGCAT  | TAGGGTATAA | AAATATGCCA  | TCAAAAACAC  | 1500 |
| 15 | AAATTGTTAA | ACATGAGAT  | ATCATTCAAG  | TAGGCAATGT | TAAATTAGAA  | GTTCTGCATA  | 1560 |
|    | CTCCAGGACA | CACGCCTGAA | AGTATTAGCT  | TTTTACTCAC | TGATTTAGGT  | GGTGGTTCAA  | 1620 |
|    | GTGTTCCGAT | GGGATTATTT | AGTGGTGACT  | TTATTTTGTG | TGGTGATATA  | GGTAGACCTG  | 1680 |
| 20 | ATTATCTAGA | AAAATCTGTT | CAATAAAAG   | GTCTACAGA  | AATTAGCGCG  | AAACAATGTT  | 1740 |
|    | ATGAGTCCTG | TCAAAATATT | AAAAATTTAC  | CAGACTATGT | TCAAACTCGG  | CCGGGTCATG  | 1800 |
|    | GTGCTGGAAG | CCCTTGTGGT | AAAGCATTAG  | GTGCCATACC | TATATCTACA  | ATAGGTTATG  | 1860 |
| 25 | AGAAAATTAA | TAACCTGGCA | TTTAATGAAA  | TTGATGAGAC | TAAATTTATT  | GAATCATTAA  | 1920 |
|    | CATCAATCA  | ACCAGCACCA | CCGCATCATT  | TTGCACAAAT | GAAACAAGTT  | AATCAGTTTG  | 1980 |
| 30 | GTATGAATTT | ATATCAATCA | TATGATGTTT  | ATCCTAGTTT | AGATAATAAG  | AGAGTAGCAT  | 2040 |
|    | TTGATCTTCG | TAGCAAAGAG | GCCTTTCACG  | GTGGCCACAC | AAAAGGAACA  | ATCAATATAC  | 2100 |
|    | CATACACAA  | AAACTTTATT | AATCAAAATTG | GTGTGTACTT | AGATTTTGAA  | AAAGATATAG  | 2160 |
| 35 | ATGTAATTGG | AGATAAATCT | ACTGTTGAGA  | AAGCGAAACA | CACTTTACAA  | TTAATTGGGT  | 2220 |
|    | TTGATAAGGT | AGCAGCTAT  | CGTTTGCCAA  | AATCAGGCAT | TTCAACCCAG  | TCCGTTTCATA | 2280 |
|    | GCGCTGATAT | GACAGGTAAA | GAAGAACATG  | TATTAGACGT | ACGTAATGAT  | GAAGAGTGGA  | 2340 |
| 40 | ATAATGGACA | CTTAGATCAA | GCAGTTAATA  | TTCCGCATGG | TAAATTATTA  | AATGAAATA   | 2400 |
|    | TTCTTTTAA  | TAAAGAGGAT | AAAATATATG  | TACATTGTCA | GTACAGGTGTT | AGAAGTTCAA  | 2460 |
|    | TTGCAGTGGG | TATATTGAA  | AGCAAAGGTT  | TTGAAAATGT | GGTGAATATT  | AGAGAAGGCT  | 2520 |
| 45 | ATCAAGATTT | TCCAGATCA  | TTAAAAATA   | TTAAGGATGT | GGAAAAATG   | AATAAGCATT  | 2580 |
|    | ATCAAAATGT | TATTATTGTT | GGCGGTACAG  | CAGGTGTTAC | CGTAGCATCA  | AGACTATTAA  | 2640 |
|    | GAAAAATCA  | AAACTTAAAA | GAGAAAAATG  | CAATTATAGA | TCCAGCAGAC  | CATCATTACT  | 2700 |
| 50 | ATCAACCAT  | ATGGACGTTG | GTTGGTGCAG  | GGGTATCTAG | TTTGAAAAGT  | TCTCGTAAAG  | 2760 |
|    | ATATGGAAG  | TGTTATACCT | GAAGGTGCTA  | ACTGGATAAA | ACAGGCTGTT  | TCAAGTTTTC  | 2820 |

55

TAGTAGCTCC AGGATTACAG ATTAATTGGT CTTCAATTAA AGGACTAAAA GAAATATAG 2940  
 GTAAAAATGG TGTITGCTCT AACTATTAC CTGACTATGT TAACgAAACT TGGAAACAAA 3000  
 5 TTTCTAATTT TAAACAAGGA AATGCCATTT TTACGCATCC AAACACTCCT ATAAAGTGTTG 3060  
 GAGGTGCGCC TATGAAATTT ATGTATTTAG CTGAAGATTA TTTTAGGAAA CATAAAATCC 3120  
 10 GT 3122

(2) INFORMATION FOR SEQ ID NO: 520:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 3982 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 520:

ATATAGATAT ATATTAATAT ATTGAGGTCA AACGATGATA ATTAAAAATT TAACAATTCT 60  
 ATTACTACTT TGTATATIGA GCTATTGGT TACAAATAGA AAGAAGCCTT TTCTGTTCTT 120  
 25 AAAGACACTC TTTATGGGTG TGGTATTAT CTTTATAGGA TATATTTCAC TGGCAATATC 180  
 TGCCGTAATT ATTTATGGTA TTATTCAATT TATCACAATT GATTTTGGTA GTTTTTCTT 240  
 AATGGGTATT ATATTGATCT TGATTTCAG TATATTCCAA TTATTTATAG TTAGATTACT 300  
 30 TTTTAGAAAA AAGAATGTCG ATTTGACAGA GGTGTGCGTT TTAsAsCATT TAATTCAATG 360  
 GTTCTTAGTT TACTTTGCGA TCTATCAAGC AGTAAATGAA AARATGGACA TTAATGATAT 420  
 TAATATCGAC AATTTCCAAT CTGTCTTTTT TGACGTGTCT AATTTGAATT TAGTAATTCT 480  
 35 ACCAACGTTA ATCATTAGCT GGGTCACAAT ATTAACTAT AGAATGAGAA GTTACAAATA 540  
 AAATCTATGA GATTATACCT TCAGACACCA ACATTCAAAT GGTGTCTTTT TTGTGTGTG 600  
 GTTTTATTTT TGAATTCGA AAAAGTAGAG GCATGAATTT TTTGACTAGT GTATAAGTGC 660  
 40 TGATGAGTCA CAAGATAGAT AGCTATATTT TGTCTATATT ATAAAGTGTT TATAGTTAAT 720  
 TAATAATTAG TTAATTTCAA AAGTTGTATA AATAGGATAA CTTAATAAAT GTAAGATAAT 780  
 45 AATTGGAGG ATAATTACAA TGAAAAATAA ATTGATAGCA AAATCTTTAT TAACATTAGC 840  
 GGCAATAGGT ATTACTACAA CTACAATTGC GTCAACAGCA GATGCGAGCG AAGGATACGG 900  
 TCCAAGAGAA AAGAACCAG TGAGTATTAA TCACAATATC GTAGAGTACA ATGATGGTAC 960  
 50 TTTTAAATAT CAATCTAGAC CAAAATTTAA CTCAACACCT AAATATATTA AATTCAAACA 1020  
 TGACTATAAT ATTTTGAAT TTAACGATGG TACATTCGAA TATGGTGCAC GTCCACAAAT 1080

EP 0 786 519 A2

|    |  |      |
|----|--|------|
|    | TCAAAATCTT GTGAGAGAAT TTGAAAAAC ACATACTGTC AGTGACACA GAAAAGCACA    | 1200 |
|    | AAAGGCAGTC AACTTAGTTT CGTTTGAATA CAAAGTGAAG AAAATGGTCT TACAAGAGCG  | 1260 |
| 5  | AATTGATAAT GTATTTAAAC AAGGATTAGT TAAATAAAC TTCAATCGTT GCTGTTATCT   | 1320 |
|    | GGAAATAATT AATTAAATGT TATGTTAATT TTTGTTAATG AAAAAGTAA TCTATTTAAT   | 1380 |
|    | GACAGGTTAA TGTAATTGTC CTGAAATTGA CTATATCTC AGTAAGTATC AATTTTAAGG   | 1440 |
| 10 | AGAGCTTATA ATGAAATTTA AAAAATATAT ATTAACAGGg ACATTAGCAT TACTTTTATC  | 1500 |
|    | ATCAACTGGG ATAGCAACTA TAGAAGGGAA TAAAGCAGAT GCAAGTAGTC TGGACAATA   | 1560 |
|    | TTTAACTGaa aGTCAGTTTC ATGATAAACG CATAGCAGAA GAATTAAGAA CTTTACTTAA  | 1620 |
| 15 | CAAATCGAAT GTATATGCAT TAGCTGCAGG AAGCTTAAAT CCATATTATA AACGTACGAT  | 1680 |
|    | TATGATGAAT GAATATAGAG CTAAGCGCGC ACTTAAGAAA AATGATTTCG TATCAATGGC  | 1740 |
| 20 | TGATGTCAAA GTTGCAATTG AAAAATATATA CAAAGAAATG GATGAAATTA TAAATAGATA | 1800 |
|    | ATAAATAAAA CAGGTTGAGA CAAAAATGG TCTTAACCTG TTTTCAATTG GCATATGTGA   | 1860 |
|    | TAAATCTAT ATCAAAATGC TTATGTATAA TGAATGACAT TTAAGAGTAG GGGAGACAAA   | 1920 |
| 25 | TATAAATACA ATAGTTCCTA GGATTACTCT CAAAATAACT ATATCAATTA TTTACTTTGC  | 1980 |
|    | TCTCCTATTT TTTAAATAT GTACATGTTT AAACAATCAA AAGTGTACAA TATTAAATTA   | 2040 |
|    | TCATTTCGAG TTCTAGTGCT ATATTGGTAG TAGTTGACTA AATGAAATA AGCTTATAAT   | 2100 |
| 30 | AAGTTTTTTC AATACTCGTG GGGCCACAAC AGAGAGAAAT AGGATCACCA ATTCCAACAG  | 2160 |
|    | ACAATGCAAG TTGGCGGGGC CCCAACATAG AGAAATGGA TCACCAATTT CAACAGACAA   | 2220 |
|    | TGCAAGTTGG GGTGGGGCCC AACACAGAAG CTGGCGAATA GTCAGCATA CAAAAATGTC   | 2280 |
| 35 | AAGTTGGCGG GGCACACA GAGAGAAATA GGATCACCAA TTCCAACAAA CAATGCAAGT    | 2340 |
|    | TGGCGGGGGC CCCAACATAG AAGCTGGCGA AAAGTCAGCT TACAAAAATG TGCAAGTTGG  | 2400 |
| 40 | CGGGGCCCCA ACATAGAGAA ATTGGAACCC CAATTCTTAC AGACAATGCA AGTTGGGGTA  | 2460 |
|    | GGACATCGAT AAAGAAATAC TTTTCTTTTA GCAATTAGTA TTTCTTATGC ATGAGCTTTA  | 2520 |
|    | CTCATGTATT CATTTTTTAA GTACaCATTa GCTACAGCTa ATGATAAAGA ACCACTACAT  | 2580 |
| 45 | AATAAATCAT TAGTGGTTCT TTATCATTTT TATCTCACTC TTTTACTGGA AGAAAAAGTT  | 2640 |
|    | TACGTTTGTa GAACATGCCA CAATACCAAA AATAATTAAg AAAAATAAGA CGATAAGCAT  | 2700 |
|    | GATGACACTT TTCAAAACA CTTCTATCAGT TTCTTTTCGAT TTTCTTTGTT GAACCTTTTT | 2760 |
| 50 | ATAACTTTCa AGTAGTTTTG CGGCTTTTTT ATTTATATGT TTATTCATGA TGTTGACTCC  | 2820 |
|    | TTATAATATA TGTTTAATTC ATTAATAATG TTGAAAACAT GACTTGAATt AAAGATATAA  | 2880 |
| 55 |  |      |

EP 0 786 519 A2

AATTGATGA TGATATTGCG TTTTATTTT CCAATGGAA TTTACTTAAA CTGATGCATT 3000  
 AAAATATTA TGAAGCACTA GAATACATAA ATGAATAGTA ATGGTGcACA GTATAGAATA 3060  
 5 ATTAAGGCTA TATTAAGTAT AATATCGTT AACTGTAAGC TATCTTTAGT TTTAATATAA 3120  
 ACTATTAGGA TAATCGACGT AAGAAGAATC ATATATATTA ATGATGAAGA AGTCATATA 3180  
 AATCCGCAT CATTTGTTGT TAATAATGGG ACTATAATTA ATCCGAAATT AATCATGCAT 3240  
 10 GCTATATATA CTATAATGTT ATACACAATG TTAATTTTTG TTCACCACCT TATACTTCTA 3300  
 TTTTAAAAAC TTCTTTATAA TGATATATTG TTTAATGTTG AAAAATTTAG ATTTACTAAT 3360  
 TTTCAITTCG TTTACATGTA AAAGGCTATA TATAGTATGC TCTTTATGAT TCTAAATGCT 3420  
 15 TTTTAATATT TAATGCTCAT CAACATTGG ATTTTGAATA TTCAATTCAA AACTTTTATT 3480  
 AGCTACGTCA ATTGTAAAT CAGAACCATA GTTGACATGA GCTACTTTTA ATTTTCCATC 3540  
 20 TAAATAATAG ATTGCGATTG CAACATCGTA AAATTCGTCA ATGACAAATA AACTCTTTTC 3600  
 GTTTGTACAC ACCTCATGCT CTCCTGAGTA TACAACGTTA ATTTCCCAAT CATTAATAAAC 3660  
 CATTTGTTAA CCTCCTTGAA CATTTAAATT GATTCAACTT AAGTTAACT TATTTCATACA 3720  
 25 ACTTCGTACA ATATCTAGAT GAACATTAAT TGTATTTCTA GAAATCTTTT TCAATTATAT 3780  
 GTACTAATTA TACTTTTAAA TTCTTATTT CAGTATAGTT TTAACACGAT TTTAAATATA 3840  
 TTCTGCAAT ATATTAACAC ATAATGTGTT CAAAAAGTTT TGAACAATTT CAAAACITTT 3900  
 30 ATATAAAGGG ATTGACAACA TGGATTCAAA TnTCTTATT TAAAAATTAC CTCATATAGT 3960  
 GTCATGTTAG CCAATTTTTA AG 3982

(2) INFORMATION FOR SEQ ID NO: 521:

35 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1353 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 40 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 521:

45 AGCTTGGATG ATTTAATaTG GtCCCTTCCC AaCCTTAGAT AATGAAAGAT tTGATAATCC 60  
 TGAATATAAA GAAGCTATGA AAAAATATCA ACAGAGATTT ATGGCTGnAG ATGAGGCTTT 120  
 GAAGAAATTT TTTAGTGAAG AGAAAAAAAT aAAAAATGGA AATACTGATA ATTTAGATTA 180  
 50 TCTAGGATTA TCATATGAAA GATATGAAAG TGTATTTAAT ACTTTGAAAA AACAAAGTGA 240  
 GGAGTTCTTA AAAGAAATTG AAGATATAAA AAAAGATAAC CCTGAATTGA AAGACTTTAA 300

55

EP 0 786 519 A2

GTTAGGTAAC ACATTTTATC AAAACTATAG AGATGATGTT GAAAGITTAT ATAGTAAGTT 420  
 AGATTTAATT ATGGGATATA AAGATGAAGA AAGAGCAAAAT AAAAAAGCAG TTAACAAAAG 480  
 5 GATGTTAGAA AATAAAAAAG AAGACTTAGA AACCATAATT GATGAATTTT TTAGTGATAT 540  
 AGATAAAACA AGACCTAATA ATATTCCTGT TTTAGAAGAT GAAAAACAAG AAGAGAAAAA 600  
 TCATAAAAT ATGGCTCAAT TAAATCTGA CACTGAAGCA GCAAAAAGTG ATGAATCAAA 660  
 10 AAGAAGCAAG AGAAGTAAAA GAAGTTTAAA TACTCAAAAT CACAAACCTG CATCTCAAGA 720  
 AGTTTCTGAA CAACAAAAAG CTGAATATGA TAAAGAGCA GAAGAAAGAA AAGCGAGATT 780  
 TTTGGATAAT CAAAAATTA AGAAAAACAC TGTAGTGCA TTAGAATATG ATTTTGAGCA 840  
 15 TAAACAACGT ATTGACAACG AAAACGACAA GAAACTGTGT GTTCTGCAC CAACAAGAA 900  
 ACCAACATCA CCGACTACAT ATACTGAAAC AAGACACAG GTACCAATGC CTACAGTTGA 960  
 CGCTCAAACT CAGCAACAAA TTATTTATAA TGCACCAAAA CAATTGGCTG GATTAAATGG 1020  
 20 TGAAGTCAT GATTTCACAA CAACGCATCA ATCACCACAA ACTTCAAAAT ACACGCATAA 1080  
 TAATGTTGTT GAATTTGAAG AAACGTCTGC TTTACCTGTT AGAAAAACAG GATCACTGGT 1140  
 25 TGGTATAAGT CAAATTGATT CTCTCATCT AACTGAACGT GAGAAGCGTG TAATTAAGCG 1200  
 TGAACACGTT AGAGAAGCTC AAAAGTTAGT TGATAATTAT AAAGATACAC ATAGTTATAA 1260  
 AGACCGAATA AATGCACAAC AAAAAAGTAA TACTTTAAGT GAAGTCATC AAAAAAGTTT 1320  
 30 TAATAAACAA ATCAATAAAG TATATAATGG GCA 1353

(2) INFORMATION FOR SEQ ID NO: 522:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1987 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 522:

GTCTGAGTCA GGTGCTGTTT GGTTAGATGC TGA AAAAACA AGTCCTTATG AATTTTATCA 60  
 45 ATTCTGGATT AATCAATCAG ACGAAGATGT AATTAAATTC TTAATAACT TTAATTTCTT 120  
 AGGAAAAGAA GAAATTGATC GCTTAGAACA ATCTAAAAAT GAAGCAGCGC ATTTACGTGA 180  
 AGCTCAAAAA ACATTAGCTG AAGAAGTAAC TAAATTTAT CATGGTGAAG ATGCATTAAA 240  
 50 TGATGCAATC CGTATTTTAC AAGCATTAT TTAGGTGAT TTAATAATCAT TATCAGCGAA 300  
 AGAATTAATA GATGGATTTA AAGATGTGCC TCAAGTGACA TTATCAAAATG ACACAACAAA 360

EP 0 786 519 A2

|    |  |      |
|----|--|------|
|    | TGTTAACAAT GGTGCGATT ATATTAATGG TGAGAGACAA CAAGATGTTA ATTATGCTTT   | 480  |
|    | AGCACCAGAA GATAAAATTG ATGGCGAATT TACGATTATT CGTCGGGTA AGAAAAATA    | 540  |
| 5  | CTTCATGGTT AACTATCAAT AAATATAATT GCATAGCTAA ATAAATTAGA GCCTACTCAT  | 600  |
|    | ATTCAATTCCT AAGAATGTAA TGAGTAGGCT CTTAATGTAC TTTTCTGTCT GTAAATTATC | 660  |
|    | TAAAGAAATC TCATCGCCT TGGCCAGGTG ATTGAAGTTC TGATCGGCTT TGACGTTTTG   | 720  |
| 10 | GTGTTCCTTT TTGTGTTTTT AATTTCACIT TAACCTCTTT TGTTTTACCA TCACGGATAA  | 780  |
|    | CGGTAAACAGT GACTGATTCA CCAGGTTTTT TATTTTCATA TAAATAGCTT CTTAATCAAA | 840  |
| 15 | CATCATCTTT AATTTTCTTG CCATCAATTT CTGTAATAAT ATCACCTTTT TTAAGATCAA  | 900  |
|    | TATCACTATC AGCTTTGGCG ACATAAATAC CGTCTTCTCT ATCAGTAGTA AGTTGCTGCG  | 960  |
|    | GCTCTTCTTC AGGAATATCT TTCAAATTAA TTAACCAAT ACCAATCGAA GGGCGGTCAA   | 1020 |
| 20 | TTTTACCATG TTTTACAAGT TGTTCAAATTG TTACTTTAAC TTCATTACTT GGAATAGCAA | 1080 |
|    | ACCCGATACC TTCAACTTGT GTCCGAGCAA TTTTCATTGA GTTAATACCA ACTAATTATC  | 1140 |
|    | CATTAAATATC TACTAATGCG CCACCTGAGT TACCTGGGTT AATAGCAGCA TCTGTTTGAA | 1200 |
| 25 | GAACGCTAAC TTTTGTATTG CCACCAGTTG TCTCAGCGTC AATCGTACGT TCGCTTGCTG  | 1260 |
|    | AAATGATACC AGATGTTACA GAGTTAGCAA ATTGTAATCC TAATGGGTTA CCCATTGCGA  | 1320 |
|    | ATACGCTATC GCCAGTTTGT ACTTTTGAAG AGTTGGCAA TTGAATCGCT TTAATACCTT   | 1380 |
| 30 | TTGTATTTTC AATTTTAAAGT ACAGCAATAT CAGTTACTGc ATCTTTACCA ACTAATTTCG | 1440 |
|    | CTTTAACTTG TTTTTATTA TGTAATTGGA CTCTAATTC ATTTGCGCCA TCAATAACAT    | 1500 |
|    | GATTGTTTGT AACGATATAA GCTGAATTGT TGTTTAATTG aTAGATAACA CCTGAACCTA  | 1560 |
| 35 | CTCCAGCTTC AGATGGTTTA GATGATTGCG CTTTAATAA GTGCTCTACA CTTGATGCTT   | 1620 |
|    | TTGCAATGTC AATAACTCCA ACAATTGTAG GGGAGACAGA TTTTATCATT TCATGAACGG  | 1680 |
| 40 | TACCGAATT CTTCCTTGA CGTCTAATT GATTGCCACC TTTATATTT GTTGCTGTAA      | 1740 |
|    | CAGTTGAACC ATCTTTATTT AAAATTGTAC TATTTAATAC TTTGCCTATA CCAAGTACTA  | 1800 |
|    | GAAAGTGACC AATAATTCCA GCAATCAATG CAACGATGAC TGThTTAAAC CATGGAAATT  | 1860 |
| 45 | TAGGCTCTCT GTATCTAGGT GTTTGGCTAT GGTGTGTGT AGAATGATCT GTATGATTAA   | 1920 |
|    | AATCTGACAT ACTTAACCTC CATTATATGA TTTATATATG CTTTAAATTAT GTCTTTThTT | 1980 |
|    | TATAATT  | 1987 |

(2) INFORMATION FOR SEQ ID NO: 523:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1351 base pairs

EP 0 786 519 A2

(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 523:

|    |  |      |
|----|--|------|
|    | TTAGAAGTCA AATCATTAnT GGCCTCnTAT CGAGTATACT ATTAACCTCA ACTATTTTGA  | 60   |
| 10 | CAATTGCATA TATTTTAATG TGGTTTAACG GCCATATGAC aCTAaCTTTG ACCTTAACGA  | 120  |
|    | CAATAATTAC AAGCTGTTTA ACCTTATTAA TATGTAGTAT TTTTATTAACT CCACTTATAC | 180  |
|    | AAAAAATTAA GCAGTTTAAAT ATAAAACTA AGCAATTTCG TAAACGAAAT TACGCAAGCA  | 240  |
| 15 | ATGATAAAAC GTTTAATTCA CCAAAAGAAA TTTATGAATT AAATCAATCT TTTAATAAAA  | 300  |
|    | TGGCTTCTGA AATTACGCAA CAAATGAATC AAATTAATC CGAACACAA GAAAAACAG     | 360  |
|    | AACTGATTCA AAACCTTAGCC CATGATTTAA AAACACCTTT AGCAAGCAAT ATTTCATATT | 420  |
| 20 | CTGAAGGACT ACGTGATGGT ATAATCACTA AGGATCATGA GATTAAAGAG TCATACGACA  | 480  |
|    | TATTAATTAA ACAAGCAAAAC AGATTATCAA CATTATTTGA TGATATGACT CATATTATCA | 540  |
|    | CTTTAAATAC AGGTAAAACA TATCCCCCAG AATTAAATACA ACTAGACCAA TTACTTGTAT | 600  |
| 25 | CAATATTGCA ACCATAITGAG CAACGTATCA AACATGAAAA CCGCACATTA GAAGTGAATT | 660  |
|    | TCGTATGCA AATTGATGCA TTTTATCAAT ATCGAAGGCC ACTTGAGCGT ATTTTAACAA   | 720  |
|    | ACTTACTTGA TAATGCGCTA AAAATTTTCAA ATGTTGGTAG TCGCATTGAT ATTAATATTA | 780  |
| 30 | GTGAAAACGA AGATCAAGAT ACTATCGACA TTGCTATTAG CGATGAAGGT ATTGGCATT   | 840  |
|    | TACCAGAAGT ACNAGAAGT ATATTGGAAC GTACTTCAG AGTAGAAAA TCTCGTAATA     | 900  |
|    | CAAAAACGGG TGGTTCITGA TTAGGCTTAT ATATAGCTAA TGAACCTCGC CAACAAAAA   | 960  |
| 35 | ACGCAAAAAAT CAGTGTAAAG AGTGATATAG ATGTAGGAAC TACGATGACT GTAACATTAC | 1020 |
|    | ACAAATTAGA CATTACGTCA TAATCCGATT TATTTATAAA ATAAATGCA AAGACTAAAA   | 1080 |
| 40 | AGAAGCTCCC ATTAATGAGG GCTTCTTTTT TTGTTTATT AGAATAAACT TTATGGGTAT   | 1140 |
|    | CCTTCTCATC ATTTTCAAGA CTTGAAGAT TTGTAGCTTG AATAATATAT TTAGGACGTG   | 1200 |
|    | CCTTAACITC ATAATATATC CTGCCAATAT ATTCACTAC AACACCAATT GAAATTAAT    | 1260 |
| 45 | GTATGCCGCC TAATAATAAA ATAGCTGCAA TCGTTGAAAA ATATCCCGGA ATATTAACAC  | 1320 |
|    | CAGATATCAT AATATTGATG GAATAAATAG A                                 | 1351 |

(2) INFORMATION FOR SEQ ID NO: 524:

50

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 433 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double

55



# EP 0 786 519 A2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 524:

|    |  |     |
|----|--|-----|
| 5  | TTGTGTGTCAG TTAAACGCAA CGTGITACCG ACTTTTCAAG TACACAA-CA TATATATGTC | 60  |
|    | TATTCCATT TTAGCCCCTG CCATTTTCAT CATTGGTGGT ATTATGTTGT TTATTTCAAC   | 120 |
|    | ATTTAATAGC TTAGATGAAA CTGCTGAAAA TAATAACAAA ATAAAGAAAC TAATGATTAA  | 180 |
| 10 | AGGACTIATC ATTATTAACA TTTCATTIAT CGTTATGATG GTTTTAACAC CATATGGTA   | 240 |
|    | CTTGATTTA ATGCTCTATC TTATTTTCTT GTTGTTTTTA TTGTGGCAAA AGGTTTATAA   | 300 |
|    | ATTTTAATAC CAAACTATT AAACACTTCT GATATTCTTA GTTCAAAATA TCAGAAAGTG   | 360 |
| 15 | TTTATAGTGT TATCTAGTTC AGATAAATAT TTCTTACTT AAAAAACGC CCTCTCTTA     | 420 |
|    | TTTTGACCCC nAT   | 433 |

(2) INFORMATION FOR SEQ ID NO: 525:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1845 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 525:

|    |   |     |
|----|---|-----|
| 30 | CCCCCATAT GATATGCTTG GCCTATTGCG TCAATCCCTT TATATTTCGG CAATAGAGAT  | 60  |
|    | GGATGTATAT TCAATATTTT ACCTTCAAAT GAAGCTAATA AGTCGGACC AATTAGACGC  | 120 |
|    | ATGTAGCCAG CTAGAATAAT CCACTCTACC TTATCTTCAT TTAATAATGT TACTAAATGT | 180 |
| 35 | TGTTCATACG CTGCTTTTGA ATCAAATGT TTTGTTTCAT TAATATAAC AGGAATATCG   | 240 |
|    | TGCTTTTTTG CTCTATCTAT ACAAAAACGA TTTTGATGAT CGGTATATAG CGCCGTAAC  | 300 |
|    | TCAATATTTT CAAGTTTTCC TGATTCAACA TGCTCAACTA TATTTCAAA GTTACTTCCT  | 360 |
| 40 | GAACCTGATG CAAAAATCGC AATTTTAAAC ATTGTTATAC CCCCACCAAT TCAATTGCAG | 420 |
|    | TTGACTCATT TTTCACAATA TGACCAATTT GATAGGCTTC CACATTTTGT TCTGCTAAAA | 480 |
|    | TCTTCAAAGC GCGTGATACA TCTTTTTCAT CAACGATAAC CGTATAGCCA ATACCCATAT | 540 |
| 45 | TAAAAATGTT ATACATTTCA TTTGTGCTA TATTGCCITG TTGTTGTAAC CAATCAAATA  | 600 |
|    | TTTTTGGTGT TGGAAATGAT GTAGTATCAA TTCTAGCAGC ATATCCGGCT GGCAATGCAC | 660 |
|    | GTGGAATATT TTCAATAAAA CCTCCACCAG TAATATGATT CATTCCTTA ATAGAAACTT  | 720 |
| 50 | CTTTTTTCAA AGCAAGTACA GGTTTGACAT ATAATTAGT TGGTTCTAAA AAGACATCGA  | 780 |

## GC CACTTGACGC A

(2) INFORMATION FOR SEQ ID NO: 526:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1141 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 526:

|            |            |            |            |            |            |     |
|------------|------------|------------|------------|------------|------------|-----|
| TCATAAATnG | TAAATGTTAA | TCGTCATATA | ATATTAAATT | ACAACACCAT | TTTGGTTATT | 60  |
| TGAAGCTTGT | GGCGCTTGT  | GTGTGCCACC | TTGATTTTGA | TTTGAGTTT  | GATCTGTAGC | 120 |
| AGGTTGTGT  | TGATTCGCG  | AATCACTGTT | ATTAGTTGAm | CTACTGTCT  | CGTGTAGATG | 180 |
| CTTATCTTTA | TCTGTCTGAT | CATTTTGTTC | TTTTCTCAAT | AAACTATTAT | CTAAGGCGGT | 240 |
| TAATGGTATT | AATGAACCAT | AATGATTAA  | GACACGTTGA | TCTAAGAAAT | CATTTTTATC | 300 |
| ATTAATAGGT | GATAATTCTA | AGTCTTTACG | AAGTAAGTTT | GCATATTTTT | GAATGCTTTT | 360 |



GATTAAGTAT ATAGAGCACT ATTTGTATT TGTTAATATT TTCACAAAA TAAAGCCTTG 720  
 ATAAATTTTAA AATATAAATT AAGCTCAATT TTTAAAAATT TATTAGCTA CAGATAACAT 780  
 5 TTTTAAAAAA GAAAGAATC AATAAATAAA ATCAACGAAC AAAAAAGTATA GAAATAAATA 840  
 GAAATAATCA TTTACTTTTC TGAAAAATTA AATTAATATT TTATTTATAA GCTGTTTTTA 900  
 AGATTTCAAG AGGAATGAAA TGTgaGAAA TTTTCAAGAT ATGCATTAC AAGTATGGCA 960  
 10 GCATTAACTT TGTGAGCAC TTTATCACCA GCAGCATTAG CGATTGATT CAAAAATAA 1020  
 CCAGCTAATT CTGATATTAA ATTTGAGGTG ACTCAAAAGA GTGATGCGT CAAAGCATT 1080  
 AAAGAAATTGC CTAAATCCGA AAATGTAAAA AATATTATC AAGATTACGC TGTACTGAT 1140  
 15 GTAAAACTG ATAAAAAGG ATTTACGCAT TATACATTGC AACCGAGTGT TGATGGTGT 1200  
 CATGCACCTG aCAAGAAGT GAAAGTACAC GCAGACAAAT CAGGAAAAGT CGTTTTAATC 1260  
 AATGGGATA CTGATGCGAA GAAAGTAAAG CCAACGAATA AAGTGACATT AAGTAAAGAT 1320  
 GACGCAGCCG ACAAGCATT TAAAGCAGTT AAGATTGATA AGAATAAAGC GAAAAATCTT 1380  
 aAAGATAAAG TCATTAAAGA rAACAAAGTT GAAATCGATG GTGACAGTAA TAAATACGTT 1440  
 25 TATAATGTTG AGTTAATTAC AGTGACACCA GAAATTTAC ATTGGAAGT TAAATTTGAT 1500  
 GCTCAAACTG GCGAAATTTT AGAAAAAATG AACTTAGTTA AAGAAGCTGC AGAAACTGGT 1560  
 AAAGG 1565

(2) INFORMATION FOR SEQ ID NO: 528:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2870 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 528:

GGATAGGTAA ATCCATTCAA AATTGGTTGT ACTCATTTTA AATCATATTT TTTAATATAG 60  
 AAACACATAA TTAAGGAGT GATGATATTA ATGAATAAAC TTCGAGACAC TACTTTTCTA 120  
 45 TCATATTTAT TTACTATCAT ATTGTGGGA TCTGCATTTC CAATGATAAA GATTGCGTTA 180  
 AATGATTTTA GTGCAGAATC ATTGTCGGCA TTTCGTTTAA TTTTAGCAAC AATAATCTCT 240  
 TTGCCGTTTG TAATTATAAA GAAATTGCCT ACCCCTGAAC TAAGAGATAT CCCTGTTATT 300  
 50 TTTATTTTAT GATTTTGTGG ATTTGTGATA TATCACACAG CTTTAAATTT TGGTGAAaCT 360  
 TTGATTAGTG CAGGTATAcC TGGrATTCyA GTCTCTACAA CGCCTATTTT TTCTAGTGCT 420

|    |  |      |
|----|--|------|
|    | GCATTTATAG GAATATCCAT TAITTCAATA AGTAAAGATG ATTACACAAC TATTAATGTA  | 540  |
| 5  | TTAGGTGTTT TTATTATTTT ACTTGATCTT TTTAGTGAAA GTTGTATTT TACTTTCCAG   | 600  |
|    | AAAAATACA TAGAAAAATA CGGCTTCATC GCTTTCACAC TATAACAAT AATGGCAAGC    | 660  |
|    | TCACCATTIA TGCTTATTTT TATTCCTGAA ATCATCAACG ATATACACGG CGCCACTTTT  | 720  |
| 10 | ACATCAATAG TATCGGTACT TTATTTAGCT ATATTCCCTA CTATAATTC ATACGTTTGT   | 780  |
|    | CTTGCTTATA TTGTGAAGTC AGTTGGTGTG TCTGATGCAA CAATGTCTCT TTATTTAAAC  | 840  |
|    | CCATATCGTT CTTTATTATT ATCTTATCTG TTATTAGACG AGCTACCAAC AACCCTTGCT  | 900  |
| 15 | ATTATAGGCG GAATTATCAC TCTACTAGGC GTTAGTTTAA GTAACTTCIT TCAAAAACA   | 960  |
|    | TAATTATTCC AAGTCCCGCA CCTCAGAATC CAAAAACATT CGAGTGATAA AATTTTAAAA  | 1020 |
|    | ATCAAAAATA TAAAAATGAT CTAATTTCTG CAAATTTACC AATATAATA CTAAATTTTG   | 1080 |
| 20 | CAATTACAAA AGGGGTATAG TCTGAGTGTA TTCTAATACG AAAGGACTTG GTGGATATGT  | 1140 |
|    | ATTACAGTTA TGGAAATTAT GAAGCATTTG CGCGCCCTAA AAAACCTGAA AATGTAGAAA  | 1200 |
|    | ACAAATCCGC TTACTTAAATC GGATCTGGTC TAGCTTCACT TGCTGCAGCT TGTTTTITTA | 1260 |
| 25 | TAAAGGATGG TCAAAATGGAA GGTTCGAAGA TTCATATTTT AGAAGAGTTA CCTAAAGCAG | 1320 |
|    | GTGGTAGTCT TGATGGTGAA AATATGCCCT TAAAAGGCTA TGTTGTCCGC GGTGGTCGTG  | 1380 |
| 30 | AAATGGAGAA CCACTTTGAA TGTTTGTGGG ACTTATTCAG ATCTATCCCT TCATTAGAAA  | 1440 |
|    | TCGATAACGC GTCTGTATTA GATGAGTTCT ATTGGCTAAA CAAAGAAGAC CCTAACTATT  | 1500 |
|    | CTCGCTGTCG TGTATTAGAG AAACAGGGTC AACGTTTAGT CACAGACGGA GACTTCACTT  | 1560 |
| 35 | TGACTAAAAC GCGGATTAAA GAAATTTTAG ATTTATGCTT AACGAATGAA GAAGATTAG   | 1620 |
|    | ATGATGTCAA AATAACAGAT GTATTTCCCG ATGACTTCTT TAATTCAAAC TTTTGGATTT  | 1680 |
|    | ACTGGAAGAC GATGTTTGCA TTTGAACCGT GGCATTCTGC AATGGAAATG CGTCGCTATC  | 1740 |
| 40 | TAAATGCGATT CGTTCATCAT ATTAGTGGTC TCGCAGACTT TTCAGCTTTA AAATTCACTA | 1800 |
|    | AATATAATCA ATATGAATCT TTAGTATTAC CTATGGTTGA ATATTTAAAA TCGCATGGGG  | 1860 |
|    | TTCAATTGGA ATACGATGTA AAAGTCGAAG ATATTAAAA ATAGTTTACG ACAAGTCAAA   | 1920 |
| 45 | AAATTCGCCG AGAAATATTA ATTGACCGTA ATGGTAATGC AGAATCTATT AAATGACTA   | 1980 |
|    | TAAACGATCT TGTCTTTGTG ACAAACGGTA GTATTACAGA AAGCTCTACT TATGGTGATA  | 2040 |
| 50 | ATGATACACC AGCGCCACCA ACTGACGAAT TAGTGGTAG TTGGACACTA TGGAAAAATT   | 2100 |
|    | TAGCGGACAA AAGTCTGAA TTTGGTAATC CTGATAAGTT TTGCCAAAT ATTCCTAAAA    | 2160 |
| 55 | AAAGTTGGTT TGTTTCAGCA ACTTCTACAA CAAACAATAA AGAGATTATC GATACAATAG  | 2220 |

EP 0 786 519 A2

TCAATGATTC TGCATGGCAA ATGAGITTTA CAATCAATCG TCAGCAACAG TTTAAAGACC 2340

AACCTGAAAA TGAAATATCT ACATGGATTT ATGCCTTATA TTCAGATGA AACGCGCATC 2400

5 ATATTAAGAAA GCCAATTACA GAATGTAGTG GTAATGAAAT ATGCCAAGAA TGGCTGTATC 2460

ACTTAGGTGT ATCAACTGAC AAAATTGAAG ACTTAGCAAA ACATGCATCT AATACGATTC 2520

CTGTTTATAT GCCATATATC ACATCTTATT TCATGACGGC TGCTATCGGC GACAGACCTT 2580

10 TAGTCGTCCC GCATCAATCT CAGAACTTAG CATTTATTGG TAACTTTGCA GAAACAGAGC 2640

GAGACACTGT ATTACAACA GAATATTCGG TTGCTACTGC CATGGAAGCT GTTTATCAAT 2700

TACTAAATAT AGATCGTGGT ATTCCAGAAG TCATCAATAG TCCATTTGAT CTTCGCGTCT 2760

15 TAATGGATGC CATATACGAA CTGAATGACC ACCAAGATTT GCGTGAGATT ACTAAAGATT 2820

CGAAAAATGCA AAAACTCGCA TTAGCAGGAT TCCTTAAAAA GATAAAAGGT 2870

20 (2) INFORMATION FOR SEQ ID NO: 529:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 3105 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

25 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 529:

30 CnGTTTAGCT GCAAAAAGTA AATAACGACA CTGTATTGG AATTTTGCAA TTGAAACAC 60

TTTTGGTGA CATTAATCA ATTTTCAGCG AGATTGAAAG CGAATACAAA ATGCTAGAG 120

AAGAAATTTT AATTTTACTA ACTTTATGGC AAAAAGGTC TATGACGCTT AAAGAAATGG 180

35 ACAGATTTGT TGAAGTTAAA CCGTATAAGC GTACGAGAAC GTATAATAAT TTAGTTGAAT 240

TAGAATGGAT TTACAAAGAG CGTCCTGTTG ACGATGAAAG AACAGTTATT ATTCATTTCA 300

40 ATGAAAAGTT ACAACAAGAG AAAGTAGAGT TGTGAAATTT CATCAGTGAT GCGATTGCAA 360

GTAGAGCAAC ACGAATGCAA AATAGTTTAA ACGCAATTAT TGCTGTGTAA GTTTAATAGC 420

ATAAAAAGAG GTTTTCATTA AGTTGAAAAC CTCCTTTTGT TGTGGCATT AATTTTTCAA 480

45 ATGTTGACTA CTCAATCCTA AATTATAAAT AGTATAGCGC aCAAAaTGCTT AAGAAATTTT 540

TTCTATGGCA CAAATGAATG GAGCATGATT ACGTTGGTTT AAAAATTGAT ATTGCAAAAC 600

TTGCGCATGC TTTTGATCCA AAGTACTCAA GTAATCAAGC AATGCATGCT TCTCAATTTG 660

50 TCCTTCGCTA TGACCATGAT ATATAACAAG TACAATAATA CCTTCAATTG ACATTAATGA 720

TAGCAATGAA TTAATAGCTT GGATTGTCGT GTCAGGCTTT GTCAGGATAG ATTTATCACC 780

55

## EP 0 786 519 A2

|    |   |      |
|----|---|------|
|    | TATATGATGT TCAATATTTT CATGTCCATC TTTTATTAAA GAAACATGAT TGAAATCCTT | 900  |
|    | AACTTTATCA CGTGTATTTT CCRAAGCTAA ATCTTGAATG TCGAAACCAT AAACATGTCC | 960  |
| 5  | TTCTGGTACT TGTTCGGCTA AAAATAAAGT GTCATTGCCG TTACCGCAAG TTGCGTCTAC | 1020 |
|    | AACAATACTT TCTGTGTTA TATGTTGTTT AATAAGCTGT TTTGAAAAAG GGAGTATACG  | 1080 |
|    | TTCTAATTTT ATTGCTTCAC CTTAGACTTG TAACGCAAGC CTTGATAAGA ATTCTACGT  | 1140 |
| 10 | GCTAATTCAG CATCGATGCC ATTTAATACT TCCATTAT TAACACTCCA CATTGGACCT   | 1200 |
|    | ACCATGATAT CTAITGGACC ATCACCAGTA ATTCGGTGAA CGATCATTTT AGGGGAATC  | 1260 |
|    | ACTTCTAATT GGTCCCAAC TAGGTTTGTG TACTCTTCTT GAGTCATAAA AGTTAATAAA  | 1320 |
| 15 | CCTTTATCGT ATTGTTTAC CATCGGTGTA CCTTTTAAAC AATGAAGTAA ATGAATTTTA  | 1380 |
|    | ATACCTTGTA CATCCATTG TGCACTTCTT TGGCAGTAGC CATCATCATG TCATAGTCTT  | 1440 |
|    | CCCCAGGTAA GCCAATTAAT ATGTGTGTAC ATACATTGAT ATTATGCTTA CGTAATTTTG | 1500 |
| 20 | CCACACATC ATAATAAGTT TTCATATCAT GGGCAGATT GATTAATCA GATGTGACT     | 1560 |
|    | GATGGATTGT TTGTAGTCTT AATTCACCC ATAAGTATGT TCGTTGATTC AAATCTGCTA  | 1620 |
| 25 | AATATTCGAC AACATCGTCT GGTAGACAGT CAGGACGCGT ACCAATAGAT AATCCACAA  | 1680 |
|    | CACCCGGTTC TTTAAGTACA GGTTCGAATT TTTCTTTTAA TACTTCAACC GGTGTCATGT | 1740 |
|    | TATTTGTAAA TGCCTGAAAA TAAGCAATAT ATTTCTCTTC GTGCCAATTC TCATGCATCT | 1800 |
| 30 | TTTCTTTAAT TTCTTTAAAT TGTACTGCGA TTGAATCTGC ACGATTACCT GCAAGTCTC  | 1860 |
|    | CGCTACCTGC AGCAGAACAA AATGTACATC CACCATGTGC TACAGTGCCA TCGCGGTTAG | 1920 |
|    | GACAGTCAAA CCCGCCATCC AATGCAACTT TAAATATTTT TTGTCCAAAT TTATTTTTTA | 1980 |
| 35 | AATGGTAATT CCATGTGTGA TAACGTTTGT TTTCAAAAGC GTATTGGAAA TGATTGCCCA | 2040 |
|    | TATGTCATTT TCCTTCTAT AAAAAAGAG TTCTAAGTAC AGATTTTAACT ATATTTTAAT  | 2100 |
| 40 | GTTATAGTGT TTATTATAGT TTGACAAAA AGAGAGAGGA ACTATGAAAT ATGAATATAC  | 2160 |
|    | CTAAATCAGT CTGGTGGCTA GTAATTGGCA TGGCGTTAAA TATTACTGGT TCCAGTTTIT | 2220 |
|    | TGTGGCCTTT AAATACAATT TATATGAAAC AAGAACTTGG AAAAAGTTTA ACTGTGCTG  | 2280 |
| 45 | GTTTAGTGCT AATGATAAAT TCATTGGCA TGGTTATTGG AAACITATTA GGTGGTTCAC  | 2340 |
|    | TATTTGTAA ATTAGTGGA TACAAGACGA TTTTAATTGG AACTTTCACCT TGCTTTTGT   | 2400 |
|    | GTACAACGCT ACTTAATTTT TTTCACGGGT GGCCTTGGTA TGCTGTATGG CTTGTAATGT | 2460 |
| 50 | TAGGGTTTGG TGGCGGAATG ATTATTCCTG CGATATACGC TATGGCTGGA CGAGTGTGGC | 2520 |
|    | CAAAATGGCG AAGACAAACG TTTAATGCGA TATACTTAGC GCAAAATATT GGTGTGGCTG | 2580 |

55

ATCTTATTAT GTATGTGTG TTTGCGCTTG TCGCGGTAA CCAATTTAAT ATTGAAATTA 2700  
 ATGCGAAAGT TAAATATCCA ACTCATTAG ATATTACTGG TAAAAAGAAT AAAGCAAGAT 2760  
 5 TTATTTTCATT AGTACTAATT TGTGCAATGT TTGCAATTG TTGGGTTGCA TATATTCAAT 2820  
 GGGAGTCTAC AATCGCTTCA TTTACACAAT CTATTAAATAT TTCAATGGCA CAATATAGTG 2880  
 10 TTTTATGGAC AATTAACGGA ATAATGATTT TAGTAGCACA ACCATTAATT AAACGGATTC 2940  
 TCTATCTGTT AAAAGGAAAC TTAAGAAGC AAATGTTGT CGGCATCATC ATTTTATGT 3000  
 TGTGTTCTTT TGTCAAGAGT TTTGCCGAAA ACTTTACAAT ATTTGyTGT GGTATGATTA 3060  
 15 TTTTAACTTT TGGAGaATGT TTGTATGGCC AGCAGTTCCA ACTAT 3105

(2) INFORMATION FOR SEQ ID NO: 530:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 5532 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 530:

TAATGATTAA ACCAGATGCA GTACAAAGAA ATCTAATTGG TGAAGTAATT TCAAGAATTG 60  
 AAAGAAAAGG ACTAAAACCT GTCCGTGGTA AATTAATGCA AGTACCAATG GAACTTGCTG 120  
 30 AAACACATTA TGGTGAACAC CAAGGTAAAC CATTTTATAA TGATTTAATT TCATTTTATTA 180  
 CATCAGCACC AGTGTTCGCA ATGGTAGTTG AAGGTGAAGA TGCAGTTAAT GTATCTAGAC 240  
 ATATTATTGG CAGCACCAAT CCTTCAGAAG CTTACCGAG ATCAATTAGA GGTGATTTAG 300  
 35 GTTTAACTGT TGGTAGAAAT ATCATTACAG GTTCAGATTC ATTAGAGTCT GCTGAACGTG 360  
 AAATTAATCT ATGGTTTAAT GAAAATGAAA TTACTAGCTA TGCTTCACCA CGTGATGCAT 420  
 40 GGTATATGTA ATAAAAATA AACTGTAAAC CTTTACGATT TATTTATAAA GGTAGAAAGG 480  
 GTTTTGTATT GTGGTTAGTC ATTATGaTTA TACATAACAA GGCCCGTTTT TTATGTTGTA 540  
 GTAAATTAAT TGAAAAATTT TATAGTTTTT kGGTAacACG TaTtAAAAg AGAGGAATAT 600  
 45 TCCTTTATCA ATGAAACTAA ACAGAGAGAA GGGGTGTGTA AAATGAAGAA TATTATTTTCG 660  
 ATTATTTkGG GGATTTTAAT GTTCTTAAAA TTAATGGAAT TACTATATGG TGCTATATTT 720  
 50 TTAGATAAAC CACTTAATCC TATAACAAAA ATTATTTTAA TACTGACTCT CATTTATATT 780  
 TTTTATGTAT TAGTAAAGAA ATTGATTATA TTTTGAAGT CAAAGATATA CAAAGCGCT 840  
 TAACATATGT ATATTTTAAAT ATCATAATTT TTTTAAACGG ACTGATTAACT TTTTAAATA 900



## EP 0 786 519 A2

|    |   |      |
|----|---|------|
|    | GATACGATTA TATTA AAAACG GCTAATCATT TTTAATTAAT GATTATATGA TGCAACTGTT | 1020 |
|    | TAGAAAATCA TGATACITTT CTACAGACGA ATATATTATA ATTAATTTTA GTTCGTTTAA   | 1080 |
| 5  | TATTAAGATA ATTCTGACAT TTAAAATGAG ATGTCATCCA TTTTCTTAAT TGAGCTTGAA   | 1140 |
|    | AACAACATT TATGAATGCA CAATGAATAT GATAAGATTA ACAACATATT ATAATGTTAT    | 1200 |
|    | CGTGGAACTA TGAAGGAGC GAGTGTGTAT GAGATACCTA ACATCAGGAG AATCACATAG    | 1260 |
| 10 | ACCTCAATTA ACAGTTATTG TTGAAGGTGT ACCTGCAAAAT ATAGAAAATTA AGGTTGAGGA | 1320 |
|    | TATTAATAAA GAAATGTTTA AGCGTCAAGG CGGTACGGA CGTGGACGTC GTATGCAAAAT   | 1380 |
|    | TGAGAAAAGT ACAGTAGAAA TAGTATCAGG CGTTAGAAAT GGTATACAT TAGGTAGTCC    | 1440 |
| 15 | AATTACTATG GTTGTAAACCA ATGATGACTT TACGCATTGG AGAAAAATTA TGGGAGCAGC  | 1500 |
|    | TCCAATAAGT GAAGAAGAAC GTGAAAATAT GAAACGTACT ATTACAAAAC CAAGACCTGG   | 1560 |
|    | TCATGCAGAT TTGGTTGGAG GTATGAAATA TAATCATCGT GATTACGAA ATGTGCTAGA    | 1620 |
|    | GCGATCATCT GCTAGAGAAA CAGCAGCTCG AGTTGCAGTC GGTGCTTTAT GTAAAGTGTT   | 1680 |
|    | ATTACAACAG TTAGATATCG ATATATACAG TCGTGTGTGT GAAATAGGTG GAATTAAAGA   | 1740 |
| 25 | TAAAGATTTT TATGATTCAG AAACATTTAA AGCAAACTCT GATCGTAATG ATGTTCTGCT   | 1800 |
|    | AATTGATGAC AGTATCGCAC AAGCAATGCG AGATAAAAT TACGAAAGCT AAAATGAAGG    | 1860 |
|    | AGATTCGAAT GCGCGTGTCG TTCAAGTTGT AGTTGAAAAT ATGCCTGTTG GTGTAGGTAG   | 1920 |
| 30 | TTATGTGCAT TATGATCGTA AGTTAGATGG TAAGATTGCA CAAGGTGTTG TCAGCATAAA   | 1980 |
|    | TGCTTTTAAA GGTGTAAGCT TTGGTGAAGG ATTTAAAGCA GCTGAAAAGC CAGGTAGTGA   | 2040 |
|    | GATTCAAGAT GAAATTTCTAT ATAATAGTGA AATTGGTTAT TATCGTGGAT CTAATCACTT  | 2100 |
| 35 | AGGTGGTTTA GAAGCGGCTA TGTCAAATGG AATGCCAATT ATCGTTAATG GTGTAATGAA   | 2160 |
|    | ACCAATTCCA ACGTTATATA AACCATTAAA TTCAGTAGAC ATTAATACTA AAGAAGACTT   | 2220 |
|    | TAAAGCAACA ATTGAACGTT CTGATAGTTG TGCTGTTCC TGCAGCAAGTA TCGTCTGCGA   | 2280 |
| 40 | ACATGTCGTA GCATTTGAAA TAGCAAAAGC ATTATTGGAA GAATTCGAAT CAATCATAT    | 2340 |
|    | TGAGCAACTT AAACAACAAA TTATTGAGCG CAGACAATTA AATATTGAGT TTTAACACA    | 2400 |
| 45 | AGAACAATTG AGGTGTAATC ATGAAATTAC AAACAACATA CCTTCAAAT AATTATCCAA    | 2460 |
|    | TATATGTGTA ACACGGTGCA ATTGACCATA TTAGCACGTA TATTGATCAG TTGATCAAA    | 2520 |
|    | GTTTATATTT AATTGACGAG CATGTAAATC AATATTTTGC TGATAAATTT GATGATATTT   | 2580 |
| 50 | TATCATATGA AAATGTACAT AAAGTTATTA TTCCAGCTGG TGAAAAGACG AAAACATTG    | 2640 |
|    | AGCAATATCA AGAAACATTA GAGTATATTT TATCCCATCA TGTAACCTGT AATACAGCAA   | 2700 |

55

EP 0 786 519 A2

|    |  |      |
|----|--|------|
|    | ACGAGGCGTG CACTTTATAC AAGTGCCAAC GACTATACTA GCGCATGATT CTAGTGITGG  | 2820 |
|    | CGGTAAAGTG GGTATTAAGT CAAAGCAAGG TAAAAACCTT ATCGGTGCAT TTTATCGTCC  | 2880 |
| 5  | AACGTGCTGTG ATTTATGATT TAGTCTTTTT AAAGACGTTA CCATTTGAGC AAATATTAAG | 2940 |
|    | TGGCTATGCA GAAGTTTATA AGCATGCGTT ATTGAATGGT GAATCAGCGA CGCAAGATAT  | 3000 |
|    | CGAACAGCAC TTTAAAGATA GAGAGATATT ACAGTCATTA AATGGTATGG ATAAATATAT  | 3060 |
| 10 | TGCTAAAGGT ATTGAACGA AGCTGGATAT TGTATTGCA GATGAAAAAG AACAAAGGTGT   | 3120 |
|    | ACGTAAATTT TTAATTTAG GTCATACATT TGGTCATGCT GTTGAATACT ATCATAAAAT   | 3180 |
| 15 | ACCTCATGGT CATGCACTGA TGGTTGGCAT TATCTATCAA TTTATAGTTG CGAATGCTTT  | 3240 |
|    | GTITGATTCT AAGCATGATA TTAATCATT TATTCAATAT TTAATACAAC TCGGCTATCC   | 3300 |
|    | TTTAGACATG ATAAGTACT TGGATTTTGA AACGTTATAC CAATATATGC TAAGTGATAA   | 3360 |
| 20 | AAAGAATGAT AAGCAAGGTG TACAAATGGT CTTGATTAGA CAATTTGGAG ATATCGTGT   | 3420 |
|    | ACAACATGTT GATCAACTAA CATTACAACA TGCAATGTGA CAATTAAAAA CATATTTTAA  | 3480 |
|    | GTAGGTGAAT GAAATGGTAA ATGAACAAAT CATTGATATT TCAGGTCCGT TAAAGGGCGA  | 3540 |
| 25 | AATGAAGTG CCGGGCGATA AGTCAATGAC ACACCGTGA ATCATGTTGG CGTCGCTAGC    | 3600 |
|    | TGAAGGTGTA TCTACTATAT ATAAGCCACT ACTTGGCGAA GATTGTGCTG GTACGATGGA  | 3660 |
|    | CATTTTCCGA CTGTAGGTG TAGAAATCAA AGAAGATGAT GAAAAATTAG TTGTGACTTC   | 3720 |
| 30 | CCCAGGATAT CAATCTTTTA ACACGCCACA TCAAGTATTG TATACAGSTA ATTCTGTGTAC | 3780 |
|    | GACAACACGA TTATTGGCAG GTTTGTTAAG TGGTTTAGGT ATTGAAAGTG TTTTGTCTGG  | 3840 |
|    | CGATGTTTCA ATTGGTAAAA GGCCAATGGA TGTGTCTTG AGACCAATTGA AACTTATGGA  | 3900 |
| 35 | TGCGAATATT GAAGGTATTG AAGATAATTA TACACCATT TATTATTAGC CATCTGTCAT   | 3960 |
|    | AAAAGGTATA AATTATCAAA TGGAAAGTGC AAGTGCACAA GTAAAAAGTG CCATTTTATT  | 4020 |
| 40 | TGCAAGTTTG TTTTCTAAGG AACCGACCAT CATTAAAGAA TTAGATGTAA GTGCAAAATCA | 4080 |
|    | TACTGAGACG ATGTTCAAAC ATTTTAATAT TCCAATTGAA GCAGAAAGGT TATCAATTAA  | 4140 |
|    | TACAACCCCT GAAGCAATTC GATACATTAA ACCTGCAGAT TTTTCATGTT CTTGGCGATAT | 4200 |
| 45 | TTCACTGCGA GGGTCTTTA TTGTTGCAGC ACTTATCACA CCAGGAAGTG ATGTAAACAAT  | 4260 |
|    | TCATAATGTT GGAATCAATC CAACACGTTT AGGTATTATT GATATTGTTG AAAAAATGGG  | 4320 |
| 50 | CGGTAATATC CAACTTTTCA ATCAAAACAAC TGGTGTGTA CCTACTGCTT CTATTCTGAT  | 4380 |
|    | TCAATACACA CCAATGCTTC AACCAATAAC AATCGAAGGA GAATTAGTTT CAAAAGCAAT  | 4440 |
|    | TGATGAAGTG CCTGTAATAG CATTACTTTG TACACAAGCA GTTGGCACGA GTACAATTAA  | 4500 |

55

EP 0 786 519 A2

AACTTGTGTTA GGGTTTGAAT TACAACCAAC TAATGATGGA TTGATTATTC ATCGTCAGA 4620  
 ATTTAAACA AATGCAACAG TTGATAGTTT AACTGATCAT CGAATAGGAA TGATGCTTGC 4680  
 5 AGTTGCTTCT CTACTTTCAA GCGAGCCTGT CAAAATCAAA CAATTTGATG CTGTAAATGT 4740  
 ATCATTTCCT GATTTTTCAC CAAACTAAA GCTTTTAGAA AATGAGGAT AATATAAAAT 4800  
 GGAAGATATC TATAAATTAA TAGACGATAT CAATCTACAA AAACTAGAAA ATTTAGACTC 4860  
 10 TGGTGTAAAT GAAGCAATAA CTACTGACAA CGATGACGCA TTATTTATTC TAGGAGAGAC 4920  
 ACTTTACAAT TTGGATTAA TGCCaCAAGG TTGGAAGTA TTCGCGTGT TATATCACAA 4980  
 ATATCCAGAC GAAAGTGAAT TGCTGATTTA TTTTATTGAA GGTTAATGT CTGAAATCA 5040  
 15 AACTGACGAA GGGTTAGAAT ATTTAAGTTA TGTGTAACCA TCACCTGAAA AGTTGATGTT 5100  
 AGAAGCAGAT TTATATCAAC AAATTAATAT GATGGAAGTT GCTATTGATA AATTACAAGA 5160  
 AGCATTGAA CTAGACCAA ATGATCCAAT AATCCATTTT GCATTGGCTG AAATGTTATA 5220  
 20 TTATGATGGT CAATATTTAC GTGCTACCTC TGAATACGAA ACCGTTTTCG AAACGCTGTA 5280  
 ATATCAAGTT AATGGTGTA ACTTATTCTC TCGTATGGCA GATTGTAGTT TACAAAGTGG 5340  
 25 KAACATAGT GATKCCGATC CgCTTATACG ATGTAATTAA TGAAGATGAA ATGACTTCAG 5400  
 AAGATTATCT CAAAGAGCC ATTTCTnACG ATAAAAATGA CATCACTCAA GAAGCAATTA 5460  
 AAATAATGAC TACATTACTT TCTAAAGATC CTGATTATAT TCAAGGCTAC TTGATTATTAC 5520  
 30 aATCaTTATA TG 5532

(2) INFORMATION FOR SEQ ID NO: 531:

(i) SEQUENCE CHARACTERISTICS:  
 35 (A) LENGTH: 942 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 531:

AATTGGGTGTTA TACTATAGGT AAATTTAAGG AGGTAAGAAA ATGGATAAAA AAGAATTAGC 60  
 45 GAAATTTATA GGCAATAAAA TCAGATACTA TAGAACCaaa TTGaACTTAA CTCAAGATCa 120  
 ACTTGAGGAA AAACCTmACA CTAAAAaGC TACTATTTCa AATTATGAGA CAGGGTACAG 180  
 AACTCCTAAA CAAGATGATT TGTTTGAAAT TGCTCATATT TTAATATACA GTATCGATGA 240  
 50 TTTGTTTCTT ACAAGAAATA ATAAAAAAA GGACATCACT TCCATATACA ACAAACCTCAC 300  
 ACCTCCCCGC CAAGAAAACG TACTTAACTA GCAAAATGAG CAATTAGATG AACAGAATAA 360

55

## EP 0 786 519 A2

AACTGGTGCT GGCATAGGAG AAGAATTATA TGAIGACATA TTGCATGAAG AAGTATTTTT 480  
 TAAAGAAGAC GAAACGCCAT CAAATGCTGA TTTTGTATT TTAGTTAATG GTGATTCAAT 540  
 5 GGAACCTATG TTAACCAAG GAACATACGC TTTTATTAAG AAAGAAGATT CTATTAAAGA 600  
 TGGTACAATT GCATCGTTG TATTAGATGG AGTAAGTCTT ATCAAGCGTG TAGATATATG 660  
 CGAAGACTAT ATTAAATTGG TATCTCTAAA TCCGAAGTAT GATGATATCA AAGTCGCTTC 720  
 10 GTTTAGTAAT ATTAAAGTAA TGGGCAAGT TGTATTGTGA TTAATAGCGC CTATATGGCA 780  
 CTTTAATATA AAAGAGTCT ATTTACGACG TGTTTAAAAG GAGTTTATAA TGAAAATAAC 840  
 TAATTGCAAA ATAAAAAAG AAACATATAGT ATATGAAGTT TTAACTAGTG GTAATCAACC 900  
 15 ATTCACATTAT GAGTTACCTA AAGATTTATC GTCACATAAT GC 942

(2) INFORMATION FOR SEQ ID NO: 532:

20 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 417 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 532:

TGGCAATGCG AAACCAAAAT ATGATCCTCG TGCAAGTTGA AGCGGGACGT TATGAAGAAT 60  
 30 GGTAAAGAA TGGTTATTTT AAACCGTCAG AAGATAAATC AAAAGAAACA TATACAATTG 120  
 TTATCCCGCC ACCAAATGTA ACTGGTAAAT TACATTTAGG ACATGCATGG GATACGACTT 180  
 35 TACAAGATAT CATTACACGT ATGAAACGTA TGCAAGGATA CGATACGTTA TACTTACCAG 240  
 GTATGGATCA TGCTGGTATT GCGACACAGG CAAAGGTAGA AGCTAAATTA AATGAACAAG 300  
 GAATAACTAG ATATGATCTT GGTGCTGAAA AGTTTTTAGA ACAGGCATGG GATTGGAAG 360  
 40 AAGAGTATGC GTCAATTTATT CGTGCGCAAT GGGCTAAAT AGGTCTAGGT TTAGATT 417

(2) INFORMATION FOR SEQ ID NO: 533:

45 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 733 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 533:

GATCCTGAAC CCGCATTTGT TTCCACTAAA ACAGTATGCC CACTTTCTAC TAAAGCGTGC 60

55

EP 0 786 519 A2

ATTTCATAC CATCCACCTC CATAATCATC TTAACGCGAA CATTITGAAA GCGCAATCAA 180  
 AAATCCACAA AATTGTAAAG GTTATTACAC TGACTTTTCC GAAAATTGTG GTAAATATATA 240  
 5 ATTAAGAAAG AACAAAGGAG CACTTACTAT GATTACTTAC AAAAATATTT TAATCGCAGT 300  
 TGACGGTTCA CATGAAGCGG AATGGGCATT TAACAGAGCA GTTGTGTGTT CTAAACGTAA 360  
 CGATGCGAAG TTAACAATTG TGAATGTAAT TGATCAAGA ACGTATTCTT CTTATGAAGT 420  
 10 TTATGATGCT CAATTTACTG AAAAATCTAA GCATTTTGCA GAAGAATTAT TAAATGGTTA 480  
 TAAAGAAGTA CCTACTAAGC CTGGTGTTAA AGATGTAGAA ACGCGTCTAG AGTTTGGSYTC 540  
 TCYTAATCT ATCATTCTTA AAAAGCTTGC ACATGAAAT AATGCAGACT TGATTATGAG 600  
 15 TGGTACATCA GGCTTAAATG CCGTGGAAG ATTATTGTT GGTTCGTAT CAGAATCTAT 660  
 CGTTCGTCAT GCGCAATGTC ACGTGTTAGT TGTTCGTACT GAAGAGTTAC CAGCAGACTT 720  
 20 CCAACCACAA GTT 733

(2) INFORMATION FOR SEQ ID NO: 534:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 6060 base pairs  
 25 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 534:

TGATGATCCT GAAGCGCTAT TGGATAATTA CAACACTGAA GATGTTGATG CACACAATTA 60  
 35 CAATAATATA AATCATGTTA TTTTGGCCTG CGATCGCGGT ATGGGTTCTA GTGCAATGGG 120  
 TGCAAGCATG TTACGTAATA AATTTAAAAA GGCGGGCATT AATGATATTA CAGTTACAAA 180  
 TACTGCGATT AATCAATTGC CAAAAGATGC TCAATTAGTT ATCACTCAGA AAAAACAATC 240  
 40 TGATCGTGCT ATTAACAAAA CACCAAATGC CATCCATATT TCAGTGGATA ATTTCTTAA 300  
 TTCACCAAGA TATGAAGAAC TTTTAAATAA TCTAAAAAAA GATGATCAAG CATAATAATT 360  
 AAATAAATTA AAAAATGGAG GATACCGCCA TGTTATTGAG TACACGTGAA AAAGAAATGA 420  
 45 TAGCCCTAT GATTAACTAC CACGGTCAAT ATATCACTAT ACACGACATT GCTCAGCAAC 480  
 TTGCGGTGTC CTCTCGTACT ATTCACCGTG AATTAAAAAG TGTTGAAGCA TATTTAACTT 540  
 CATTTCATT AACTTTAGAA CGCGAAACA AAAAAGGGGT ACGCATTGCT GGCACAGATT 600  
 50 CTGATTTAAA CGATTGGAAG CAATCGATTG CACAACATCA AACCATTGAC TTATCTGTTC 660  
 AAGAGCAGAA AGTAATTATT ATATACGCTT TGATACAAGC CAAGGAGCCA GTTAAACAAT 720

55



EP 0 786 519 A2

|    |   |      |
|----|---|------|
|    | TAACAGCCAA AATGAAGCAA TTGAAAAAGC AGGTAAAGCC TTAGTTGATA GTGGTGCTGT   | 2640 |
|    | AACAGATGCT TATATTCAAG CAATGAAAGA TCGTGAGCAA GTCGTATCAA CATTTATGGG   | 2700 |
| 5  | AAATGGCTTA GCAATTCCCTC ATGGCACAGA TGAAGCTAAA ACAAATGTGA TTCACCTCAGG | 2760 |
|    | TTTAACATTAA TTACAAATCC CTGAAGGCGT TGACTGGGAT GGCAGAGTAG TTAAGTTGT   | 2820 |
| 10 | CGTGGGAATT GCTGGTAAAG ATGGCGAACA TTTAGACTTG TTAICTAAAA TTGCAATTAC   | 2880 |
|    | ATTTAGCGAA GAAGAAAAATG TGGATCGTAT CGTTCAGCA AAATCTGCAG AAGAATTTAA   | 2940 |
|    | ACAAGTATTC GAGGAGGCGAG ATGCATAATG AAAGCAGTTC ACTTTGGTGC TGGTAACATA  | 3000 |
| 15 | GGTCGTGGTT TCATTGGTTA TATTCTgCAG ACAACAATGT TAAAGTAACA TTGCAGACG    | 3060 |
|    | TCAATGAAGA AATCATTAAAT GCTTTAGCTC ATGATCATCA ATACGATGTT ATTTTAGCTG  | 3120 |
|    | ATGAGTCTAA AACACGACG CGCGTGAATa ATGtGTATGC AATTAATTCA ATGCAACCTT    | 3180 |
| 20 | CTGAAGCGTT GAAACAAGCA ATTCTAGAAG CTGATATTAT TACAACAGCT GTTGGTGTTA   | 3240 |
|    | ACATACTACC TATTATTGCT AAATCTTTTG CGCCTTTCTT AAAAGAAAAA ACAACCATG    | 3300 |
|    | TTAATATTGT TGCTGTGTAG AATGCTAITA TGGCAACTGA TACATTGAAA AAAGCAGTAC   | 3360 |
| 25 | TTGATATTAC TGGCCCTCTT GGTAAChaTA TTCATTTTGC TAACTCAGCA GTTGATAGAA   | 3420 |
|    | TTGTACCATT ACAAAGAAT GAAAATATAT TAGACGTTAT GGTGAGCCA TTTTACGAAT     | 3480 |
|    | GGTGTGTGA AAAAGATGCA TGGTATGGTC CAGAACTAAA CCATATTAAA TATGTGTATG    | 3540 |
| 30 | ATTTAACACC ATATATTGAG CGTAAATTAT TAACTGTGAA TACAGGACAT GCATATTTAG   | 3600 |
|    | CGTATgCTGG TAAATTTGCA GGTAAAGCTA CAGTTTtAGA TGCAGTtGAA GATAGTTCAA   | 3660 |
| 35 | TTGAAGCTGG CTTACGCCGT GTTTTAGCTG AAACTAGTCA ATATATTACT AATGAATTtG   | 3720 |
|    | ATTTTACTGA AGCGGAACAA GCTGGTTATG TTGAAAAAAT AATAGATCGT TTCAACAATT   | 3780 |
|    | CTTATTtATC TGATGAAGTA ACAOGTGTG GACGAGGTAC ATTACGTAAA ATTGGCCCTA    | 3840 |
| 40 | AAGATAGAAT TATAAAACCA TTAACATATC TTTATAATAA AGATTTAGAA CGCAGTGTT    | 3900 |
|    | TATTAAATAC AGCTGCATTG TTATTGAAGT ATGATGATAC AGCAGACCAA GAAACTGTTG   | 3960 |
|    | AGAAAAATAA TTACATTAAA GAACACGTTT TAAAGCGTT TTTAAGTGAA TATGCTAAAG    | 4020 |
| 45 | TTGACGATGG CTTAGCCGAT GAAATAATTG AAGCGTACAA TTCACTTTCA TAAATTtATTG  | 4080 |
|    | AGCTTTGTTT GAAACAAGAA GTTCCAACG TTATTCGTTA ACAATCAGTA ATAATGTAGT    | 4140 |
|    | AGTTCCTCTG AATTAACAAT ATTAAATTTC TGAACATAAA AAATACTCCC TTCAACATAG   | 4200 |
| 50 | ACACTTAACT TGTGTTATGT ATGAAAGGAG TATTTTTGCG TTAATAATTT GTTTTATTTT   | 4260 |
|    | CGAGCCACAG CCACCTATTc AATGGCTATT GGTCAATTACT AAAACAaATT CATATTAACT  | 4320 |
| 55 |   |      |

EP 0 786 519 A2

|    |            |             |             |             |            |             |      |
|----|------------|-------------|-------------|-------------|------------|-------------|------|
|    | TTGAATAAAT | TTTTCTCTTC  | AGTTTGTGTG  | TCTTCTCTAG  | TGAATCTCTT | AATTAGAAT   | 4440 |
|    | GCCATACCTG | CACCTAGAGC  | TAATTACAGA  | TATGTTAAAT  | CGTCATTATG | TGACATACCA  | 4500 |
| 5  | GTATCTGGTA | AAGTTTATAG  | TGTGTGTTTA  | GCTTTATTAA  | CTTTTCTCTG | TTGAGCTGAT  | 4560 |
|    | TTTGTCTTAG | CTTGGTGGTC  | GTCAGTGTTA  | GTTACATTAA  | GCATATCTTG | ATTAGCACTA  | 4620 |
|    | TTGCTTCCAT | TTGAAACTGT  | AGCTGGAGAT  | GCATTGGCAC  | CGTCGTTTTG | CGTAGyTTTA  | 4680 |
| 10 | TTGTTTGCAG | CTGAACAAC   | TGATTTTTCG  | GTATCATTAG  | TATCTGCTGT | TGCCGCTATCA | 4740 |
|    | CTTTTTTTGC | TAACTATTGT  | TGAAGTCATT  | TTTCTTTTTC  | CTTCAGAGA  | TGCAGTGTIT  | 4800 |
| 15 | GATGTTTAT  | TCGAAACTTC  | AGTATCAGCT  | TTTGTGGGCG  | ATTATTCTGC | TGCTTTAGAT  | 4860 |
|    | GCAACGTTAG | TTTCAGACTT  | AAGTTGCTCT  | GCATCAGTIT  | GATTTGTGCT | ACTTCTCTCT  | 4920 |
|    | TTATCTTTTG | ATGTATTAGA  | AGGTACATTT  | GGTTCCTGTA  | TGTCGTCTGA | AGGCAATGTT  | 4980 |
| 20 | TCAGTTGTG  | ATTCAACCAT  | ACTTTGATTT  | GTGGAATCAC  | TACCATCTTT | TTCTGCCTTA  | 5040 |
|    | GCTTTATTTT | CAGATTTTGG  | TTGTGCAACC  | TTGTCAATTAG | TTGATTGAGA | TTCAAGCACTA | 5100 |
|    | TTATTACTAT | CAGCACTTTG  | TTTGTGAATCA | TTTACAGATG  | CATTATCTTT | GCTATCAGCA  | 5160 |
| 25 | GATGATGCTG | CTTCTGTGCT  | GCGAGTTGTT  | GGAGCCGTTG  | CTGTTGATCC | TGTTGGTGCA  | 5220 |
|    | TTCTCGTTTG | TTGCTGTAGT  | TGTACTATTG  | TTATTTGTTG  | TGCTTCTCTG | TGCGCCTTGCA | 5280 |
|    | TTATCAGTTT | CTGTTACAGG  | TTTATCAGTT  | TGCGCGTAT   | TAGTTGATCT | TACTTTCTGT  | 5340 |
| 30 | TTACTAGTTA | CATCGTTATC  | CATTGTCCGA  | CTGTTTGTGT  | ATGCACTTAC | ACTGAATTG   | 5400 |
|    | TTATTAGCTT | GCGGTTTATC  | ATTTGCAATCA | TCAAGTTCGT  | ATGTTGCTGT | TGTTTCAACT  | 5460 |
| 35 | GTTGCCGACT | CACATTATTAT | TGTTGTGTGC  | GGAGAAGCGT  | CTGCTTTGCC | ATTAGCTGTC  | 5520 |
|    | GTCTCAGATA | CGTTAGGTTG  | TCCAGTATTT  | TCTGTTGTTG  | CATTAGCAAT | TGAATTTGCT  | 5580 |
|    | GTTGCATCAT | TATTATCTAT  | ACCATTATTA  | GTATCAATTAG | CATCTGGAGC | ATTCTGAGGC  | 5640 |
| 40 | ACAATGCGCT | CAATTGCAGG  | TATCGTTTACA | TTTTTGTAAT  | CAGCAACTTC | TGCATTTGTT  | 5700 |
|    | TGTGTTTAT  | CTAATTTATC  | AGCAAATCTG  | TCAAATATTC  | TACCTAAATC | CGTAGCGTCA  | 5760 |
|    | ATTCTTTTGT | CCGATGATCA  | TGCATCTGCA  | TTTTTAATTA  | TTTCTATTTG | CTTGTTAAAC  | 5820 |
| 45 | ACTTCTCTGA | TTGCTTCCAA  | AGCAATTTTC  | TTAACTTCAG  | GATTAAATAG | TTGTGCTTTA  | 5880 |
|    | AGTTGTTCAA | GCGCACTATT  | TTTGACAGTA  | GCGAATTTCTG | CATTGTTAGT | TTGATCAGAA  | 5940 |
|    | ATATCTTCAG | TGTGTTTTGA  | TAAATGTGCT  | CTTAAAGCAT  | TCGTTAAACG | TTCTTTTTCT  | 6000 |
| 50 | TCAGTTGTAG | CATCAGCGTT  | GACATTTTACA | CCTGCTTCAA  | TCGTGCTAG  | TGAGCTTTCT  | 6060 |

(2) INFORMATION FOR SEO ID NO: 535:



(A) LENGTH: 977 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 535:

|    |  |     |
|----|--|-----|
| 10 | AACAAGCCT TCCAATTATC TGCCTCGGTA GAACAAGTAT TAGCAACTTT ATCACCTACG   | 60  |
|    | CTAAACAGTC CTTAGCATT TACGGCACG ACAAAAATGC TAGATATTAC ATTGATTCA     | 120 |
|    | TTGAACATG ATGGTACAAC GTACCTGTG GACTATGCTA CGTTGAAAA TGATTATGAA     | 180 |
| 15 | GATAATAAG ATCCTGAGTT TAGACGTAAA AGTTTCAAAT CGTTTAGCGA TGGGATTGGA   | 240 |
|    | AAATATCAGC ATACTACCGC GGCTACATAT AATATGCAAG TACAACAAGA AAAAATTGAA  | 300 |
|    | GCTGATTTAC GTGGATTGGA ATCAGTCATC GATTATTTAT TACATAGTCA AGAAGTAACG  | 360 |
| 20 | CGTGATATGT TTGACCGTCA AATCGATATG ATTATGCGTG ACTTGGCACC AGTATGCAG   | 420 |
|    | AAATATGCTA AACCTTTTACA ACGTATTCAC GGATTAGATA ACATGCGTTT TGAAGACTTG | 480 |
|    | AAGATTCTTG TAGACCTCTGA TTATGAACCA GAGATTTCAA TTGAAGACTC AAAAAATTAT | 540 |
| 25 | ATTTTCGGTG CGTTAAGTGT TTAGGGTAT GACTATACAA ACATGTTACG TGAAGCATAC   | 600 |
|    | GATCAGCGAT GGAATTGATT TGCACAAAAT AAAGGTAAAG ATACAGGCGC ATTTTGTGCA  | 660 |
| 30 | AGTCCATACT TTACCATTC ATATGTGTTT ATTTCTTGA CTGTAATAAT GGCTGAAGCA    | 720 |
|    | TTTGCTTAG CACATGAATT AGGTCATGCA GGTCATTTTA CATTAGCTCA AAAACATCAA   | 780 |
|    | CCATATCTTG AATCAGAAGC ATCAATGTAC TTGTTGAAG CCCCTTCTAC AATGAATGAA   | 840 |
| 35 | ATGTTGATGG CCAATTATTT ATTTAACACA AGTGATAATC CAAGATTTAA GCGTGGGT    | 900 |
|    | ATTGGCTCAA TTTTATCTAG AACATATTAT CATAATATGG TACCCmTTTA TTGAAGCnG   | 960 |
|    | CTTATCCACG GGGAGTG   | 977 |

## (2) INFORMATION FOR SEQ ID NO: 536:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1440 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 536:

|  |     |
|--|-----|
| AGACAGTGAT yGaATTTCAT TTACAGTACA CAAATCATCG AAAAAATTGGT AACATTCTTC | 60  |
| TCTATTTTCT AACGTTAgwA TTgCATCAAA CAATTCTCTC AACGCTGCAC CTCGTAATTT  | 120 |

EP 0 786 519 A2

|    |             |             |             |             |             |             |      |
|----|-------------|-------------|-------------|-------------|-------------|-------------|------|
| 5  | AATCATATAA  | AGTGTAGCGA  | TTTATATTTT  | ATTAAATCTG  | AATCGTTACT  | TTATTTTAAT  | 240  |
|    | TTATGCTAAT  | CCAGCGGGTT  | CGAAAATAGT  | GTCAACTTGA  | TTCAAATGAT  | GTTTAGGATC  | 300  |
|    | GAAACATTCA  | TCCAATCTTT  | CTTTTGTATA  | AACACTTGTA  | ATAGACTACT  | CTTGTTCTGAT | 360  |
|    | TAATTCACGG  | AACGGTGTTT  | TGTTTCCCA   | AGATATCATC  | GCTTTTGGTT  | GTACTTTTGT  | 420  |
| 10 | GTATGCTTCT  | TCAAGAACCA  | TACCTTTATT  | AATTAATGCT  | AATAAGACAC  | GTTGTGAGAA  | 480  |
|    | AATCAGACCA  | AATGTTTTAT  | CTATGTTATT  | ACGCATATTA  | TCTTCAAATA  | CAGTTAAACG  | 540  |
| 15 | GTCCACAATA  | TTTGTGAACG  | AATCAATGCA  | TAATCTAGTG  | CTAATTGTAAC | ATCTGGTAAAC | 600  |
|    | ATAATACGCT  | CAGCAGAAGA  | ATTGAGAAATA | TsTCTTTCAAT | GCCATAATGC  | CACATTTCTCA | 660  |
|    | TAAAGCTGTAG | TAAATATGAC  | ACGAATGACT  | CTTGAAATAC  | CTGTGATATT  | TGTGACACCA  | 720  |
|    | ATTGAGATTTC | GTTTATGAGG  | CATTGCAGAT  | GAACCTTTTT  | GGCCTTTTGC  | AAATGCTTCT  | 780  |
| 20 | TCAACTTCTC  | TCGTTTCGGT  | TTTTTGAAGG  | TTACGTATTT  | CAACGGCAAA  | TTTTTCTAGT  | 840  |
|    | GATGTCGCGA  | TTAATGCTAA  | TGTCGCAATA  | TAGTATGCAT  | GTGATCGCG   | TTGCAATGTT  | 900  |
| 25 | TGCGTTGATA  | CAGGCGCTGT  | GCCAATACTT  | AAATGTTTAC  | ACACATAACT  | TTCTATTTC   | 960  |
|    | GGAGGAATGT  | TAGCAAAAGT  | ACCTACTGCA  | CCACTCATT   | TCCCTACTTC  | AATTTCTTCT  | 1020 |
|    | CTTACTTGGT  | TGAAACGTTG  | TAAGTATACG  | TGCATTTCGG  | TGTACACAAA  | TGCCATTTTG  | 1080 |
|    | ACACCAAATG  | TAGTTGGTGT  | TGCATGCACT  | CCATGTGTAC  | GTCCCATCAT  | CAATGTATAT  | 1140 |
| 30 | TTATAATTTT  | TTGCTTTTTT  | AGCTAAAACG  | TCGATAAAAT  | TTTCTAAATC  | TTTTTCAATA  | 1200 |
|    | ATGTCATTGG  | CTTGTTTAAAT | AACGAAACTT  | AAAGCTGTAT  | CTACAACATC  | AGTAGAAGTT  | 1260 |
| 35 | AAACATAATG  | TGACACCTTA  | CGTTCTTCAC  | CTAGCGTTTG  | AGAAACTTGT  | CTAGTAAAGG  | 1320 |
|    | CTACAACATC  | ATGGCGCGTT  | TCTTGTTCAA  | TTTCTGTGTC  | ACGTTTCGACA | TTTACCTTTG  | 1380 |
|    | CGTTTGTGCG  | AATTTTTTGT  | ACGTGACGTT  | TGCGTATATG  | TCTTATTTCA  | CTCCATGCTT  | 1440 |

(2) INFORMATION FOR SEQ ID NO: 537:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 784 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 537:

GATAAATCTA TnCAGTTTCC GTCCAAATAT CtGCaCCTAA AGCTTTTAAg TGTCTACAA 60  
TATCTgTATA ACCTCTATAA ATATGTTTAA CATTGTAAAT TGTAGTTACA CCCTCAGCAA 120

## EP 0 786 519 A2

CATGTAATGT TGATGGTTTT ATCGTTGCTG TGCCTTCGTC AACTCAATA TTTGCACCCA 240  
 TGGGCTTTAA TTCTTCAACA TGTTTAAAC GCTCCGATA AATCGTATCA GTTACAAATG 300  
 5 AAGGACCATT TGCCATAAAT AATAATGGTG TAATAGGCTG TTGCAATCA GTAGCAAAAC 360  
 CTGGATATAC TAGTGTTTIA ATATCAACAA ATTGATATGG CGCATTATTA TTGATGCGAA 420  
 10 TTCTTTGCTC TCTTACATCA ACATTACAC CTAATTCACT AAATTTAGCA GTTAATGTTT 480  
 CTACATGTTT CGGAACAATA TTATTTAATA TAACATTTTC TCCCAATGCT GCAGCGATAC 540  
 ACATATATGT GCGTGCCTCA ATTCTATCAG GTATAACTTG ATACTCAGAA CCAATGTAAT 600  
 15 CTTTGACGCC ATTGATTITA ATTGTTGATG TACCCGCTCC CTTAATATTA GCTCCCATAC 660  
 TTGTTAAGAA GTTAGCAACA TCAACTACTT CGGTTCTTT AGCAGCATTT TCAATTACAG 720  
 TTTGTCCTGT TGCATAAACT GCAGCTAGCA TAATGtAAAT TGTTGCACCT ACGCAACCA 780  
 20 TATC 784

## (2) INFORMATION FOR SEQ ID NO: 538:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 3733 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 538:

CAATCTCCAC CAATGACAG TCGTCATTCTG TGACCTCATA CCAACAAAA AACAGTCTCG 60  
 CAATCAAGAC TGTTTTCCAC TCAATATATT CATCCATTAG CGTAATAGAT TATTTGACTT 120  
 35 CTGTAGCTAC AAAGATTITA CGTTTTTCCC AAACGCCCTGT eTTTTCATTG TAATCATCAC 180  
 AAGTAATTAA TGTTAATTGT TTATCTTTAC CTTTTTGTC ATCTAGAACT CCTACATCTG 240  
 40 TAGGCITAAC ATCTCTTATA CTGTCTATT TATACITACG TGTTTCATTA CCAACTTTAA 300  
 AGTACACCAT ACTACCTTTT TTGGCTGCTT TAAGATTGT AAATTGATAG TTGCGACGGT 360  
 CAATGAAAGT GTGCTCTGCA ATTGAAATAT TTTGATCATC TAGTGATTCA TTTTCTCTCG 420  
 45 CAAAGCTTAC ACCCTCATTT AATTGTTTCA GGTGTGCTGG TCCTGGATAT ACTGGTTCTT 480  
 TAATATCAGC ATCTGGAATT TCAATATAGC CTGCCACTTT CGATTATCTT TTGCGAATTT 540  
 GAGGTTTAGC TTGCTGCTTT TTATCTTTAC TCGCCTGTTC TTTTACATTI TTATCATATT 600  
 50 GTTCAATCTT TTCATCTTIA TCTTTATCGT GAAGATAATT ATCGATATGT GGTTTAGCAA 660  
 ACAAAATATGC TGCCACTAGG aTAAGTACCA CACCAGCGAT TGTCATTAAAT CGATTGTGTC 720

55





AAGAGATAAG AAAAAAGATG ATTGTAGATG CCACGCAAAA GCCTATTGGA GAxACTTTGC 300  
 CAATAAATCG TGGTGATTTA TCCCGATTTA TTAATAAAAA TTGGAGGAT TTTAAAAATGC 360  
 5 CAAAACGTAC TTTCACATAA GACGACATTC GTAAATTTGC AGAaGAGGAA AaTGTAaGaT 420  
 ATTTAAGATT ACAATTCACT GATATTTTAG GAACAATTAA AAATGTTGAA GTGCCTGTAA 480  
 10 GCCAATTAGA AAAAGTACTT GATAACGAAA TGATGTTTGA CGGTA 525

(2) INFORMATION FOR SEQ ID NO: 540:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1408 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 540:

TTGATTTCGT ACAAGTATC TnCTCATTTC TGATCTCTGA AAAATCTTTA GTGTAATAAT 60  
 GTTGTCAGT TTTAATATTT TCAGTCATAG TTGACTACCT CCGTATATTT TGATTTAATT 120  
 25 AAGTTGTATA TTTTGATGAA CACTTATTGT TACTTGTTGG CGCAAGTAGC AGTTTTTTCA 180  
 TTCTTCATAA AAGTATTCCT TATAGAATAT GAATGTTGCG ATACTTGCGA ATCCTGCAAT 240  
 TGACCATGCT GtAGTGAAGT ATAGAAACGG CATAAGTACA ATCGCTAAGA CTGTGAAGCA 300  
 30 TAGTACTGCT ACTAGGTAGC TTTTATAAAT GTTACTCATT TTCTTTTTTC AACTCTCCCA 360  
 TTATCTCTCG GTCTGATAAG TCGTGATAAG GGAATTTTTT CClAGCTAAT TGGACTGGTA 420  
 TTCTGCCTCG TATGCAATG TATCCTTCAT CTTCAAGCTC TTTATTCAGT TCTCTTATTA 480  
 35 TTGTCTCGC TTTGGAITTA GAAACAGATA AAATTACCgC AAGTTCITTA GCTTGCAAAC 540  
 TATTTTTCAT CATACTTTT CCTCCTTTAA AATAACIGTT GATTCTCTGG GTTATCTGCT 600  
 TCGTAATTAT CTGCAATAAT ACTTTTAGCG AAAAAAGTCCA AACTGACCTT ATATAGGTTG 660  
 40 TTCATAGATT TCTTTACGTT AACCCCTTCC TCAAGTACAT AAGGCACCTT AAAATCATT 720  
 ATAAACAGTC CGTTTTGCTC TAAAGTAACG GTTGGTAATT CAGGTTGTTT CCGTCTATAA 780  
 45 ACTTCTCCTA GTGTAGGTTT TTGCTTTTCA GCTTGTTTAG TGAAGTCGGA AAATGCCTTA 840  
 AGTAGTTTTA TTCTGAATC AGGATCACTG TGTCGCTCAA TCGTTTCTGC TGTAGACTCT 900  
 TTACTAAAAT CATTTCTATT GATTACAGGC TTTCCTGTAT TTGTTTCAAT CTTCCAAACC 960  
 50 TTCCACGTCA CAACTGCCAT TGTGGTGAGG AGGGTTGTTT TGTATAGTGC GTTCATTGTT 1020  
 AATTCTCTCT ATTAAGTTGT TTGTTCAATT GTGTGTGTTA TTCTTCTTCG TCTAAATCAA 1080

CGACTTCGTA AGTGTGCTCA ATCTGCGCTG cATATGTCAC AGTAAGAGTA TCTTTGTGTG 1200  
 TGTATGTTTG ACTTTTGTty TCTtTAACTG CATAAAGTGT TAATACTATA TTGTTTAGCT 1260  
 5 TTYCTTTTGT TTCTGGTGTC ATTTACGCTC CCCCTAmATT AGCyTCATAA CCGAATTCAG 1320  
 TCATGATTC ATGTATTTTC AATCTGCCTT TTGTGTCCA TCTAGTTTGT AAAACTGTGT 1380  
 10 CTTCCTGCGC ATCAGAACGC ACAATTGT 1408

## (2) INFORMATION FOR SEQ ID NO: 541:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 432 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 541:

GTTCGTCAAT GATTTTATC CGATGCTATG AGCATTatCA AATACAAAT GCTCTTTAA 60  
 AAGCAGTTAT TGACTGAAAA TCTACTTCTA AGAGAGCACT TTATTTAATT ACTTAAGAAA 120  
 25 TCTTGAAATT TCAATATACG ATGTTTATGA TAAGTCGCTT ATTTCATCTT TAGGCTTGT 180  
 ATTAGTAAGT AGTTTAATAC CACTGATTAA CCATAAGCA AATGTAATTA TGTTACCACT 240  
 TATTACAGCT CCAATAATCA ACAATATACC ACTCATTTTT TTGTTTTTAG ATGCTTTAAA 300  
 CATACCGATT GCACCTAAAA TAATTGAAAT GATTCCAAAT ATGAATAGGG ATAAGAATAA 360  
 TACAGTGAAA ATTGCTGCTG CTGTTTCTGc ATCAACTGGG nCAACCTCAC CATTAACGTG 420  
 35 TGTTGGACAC AT 432

## (2) INFORMATION FOR SEQ ID NO: 542:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 2426 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 542:

ATAATCATGA AGThGCTAAA nGCCAAATA ATGATGGTTC TGGACATGTT GTGTTAAATA 60  
 AATTCCTTTC AAATGAAGAG AATCAAAGCC ATAGTAATCG ACTCACTGAT AAATTACArG 120  
 50 GAAGCGATAA AATTATCAT GCTATGATTG AAAAATTAGC TAAAGTAAT GCCTCAACGC 180  
 AACATTACAC ATATCATAAA CTGAATACGT TACAATCTTT AGATCAACGT ATTGCAARTA 240

EP 0 786 519 A2

|    |  |      |
|----|--|------|
|    | GTATAAAAG TCAACGAAAT ATTATTTTGG AAGAACTGCG ACGTACTGAT GATAAAAAGT   | 360  |
|    | ATGCTACACA AAGCATTTTA GAAAGTATAT TTAATAAAGA CGAGGCAGTT AAAATCTTAA  | 420  |
| 5  | AAGATATACG TGTGTAGGT AAAACAGATC AACAAATTGC AGATCAAAAT ACTCGTCATA   | 480  |
|    | TTGATCAATT ATCTCTGACA ACGAGTGATG ATTTATTAACT GTCATTGATT GATCAATCAC | 540  |
|    | AAGATAAGTC GCTATTGATT TCTCAAAATT TACAAACGAA ATTGAGAAA GCTGAAGCAG   | 600  |
| 10 | ATAAATGGCG TAAAGATTGG ACGAATAAAG GATTATCAAA TCGCCAAATC GTTGACCAAT  | 660  |
|    | TGAAGAAACA TTTTGCATCA ACTGGCGACA CGCTTCAGA TGATATATTA AAAGCAATTT   | 720  |
|    | TGATAAATGC CAAAGATAAA AAACAAGCAA TTGAACGAT TTTAGCAACA CGTATAGAAA   | 780  |
| 15 | GACAAAAGGC AAAATTACTG GCAGATTTAA TTACTAAAAT AGTAACAGAT CAAAATAAAA  | 840  |
|    | TTTTTAATTT AGTTAAATCG GCATTGAATG GTAAAGCGGA TGATTTATTG AATTTACAAA  | 900  |
|    | AGAGACTCAA TCAACGAAA AAAGATATAG ATTATATTTT ATCACCAATA GTAAATCGTC   | 960  |
|    | CAAGTTTACT AGATCGATTG AATAAAAAAT GGAAAACGAC AGATTTAAAT AAGTTAGCAA  | 1020 |
|    | ATTTAATGAA TCAAGGATCA GATTTATTAG ACAGTATTCC AGATATACCC ACACCAAAGC  | 1080 |
| 25 | CAGAAAACGT TAACACTTGG TAAAGGTAAT GGATTGTAA GTGGATTATT AAATGCTGAT   | 1140 |
|    | GGTAATGTAT CTTTGCTTAA AGCGGGGGAA ACGATAAAAG AACATTGGTT CCGCATATCT  | 1200 |
|    | GTAAATGTGG GTGCAATGGG TGTACTAATG ATTTGGTTAT CACGACGCAA TAAGTTGAAA  | 1260 |
| 30 | AATAAGCAT AATTATATTG GGGGAAGAGC ATCTATATAT TTTTTTAAGT ATATAAGACG   | 1320 |
|    | TCTTATTTCC CCTTAATTAA TTGTGAAGTA TAGCAAAAT GCAATGAATA GATTGTCAT    | 1380 |
|    | CATTTTAAAG TTATAATGAA TTTAACGACT TAGAATCTAC CAAGTAAAGG AGAATGAAGA  | 1440 |
| 35 | TGCTCGAAAA AACGGCGCTA TTAGTTTGG ATATGCAAGA AGGTATAGCG AGTAGTGATC   | 1500 |
|    | CTAGAATAAA AAATATTATT AAAGCGAATC AGAGAGCAAT TGAAGCAGCA AGACAACATC  | 1560 |
|    | GAATACCAAG CATTTTCATA CGTTTAGTGT TAGATAAGCA TTTTAAATGAT GTCTCCTCGA | 1620 |
|    | GTAATAAAGT GTTTTCAACA ATTAAGCTC AAGGATATGC GATTACTGAA GCAGATGCAT   | 1680 |
|    | CTACACGAAT ACTTGAAGAT TTAGCACCAC TAGAAGATGA GCCGATTATT TCTAAGCGAC  | 1740 |
| 45 | GCTTTAGCGC ATTTACAGGT AGTTACTTGG AAGTTTATTT ACGTGCAAAT GATATTAAATC | 1800 |
|    | ATTTAGTATT AACGGGTGTC TCTACAAGTG GAGCTGTATT GAGCACGGCA TTAGAAAGTG  | 1860 |
|    | TAGATAAAGA CTATTATATT ACTGTTTTAG AAGATGCTGT TGGTGATAGA TCAGATGATA  | 1920 |
| 50 | AACATGACTT TATTATTGAA CAAATTTTAT CACGCTCATG TGACATTGAA TCCGTAGAGT  | 1980 |
|    | CATGGAAAAG TAGTTTATAG TTAATATAAC GTCAATTAAA GCTCGGCGAT AATGTTTGAG  | 2040 |

55



GAGGAACATT TGAACATAAA ATAATATATT TATATAAAAC GACCAGGCG TTCGAAGTGA 2160  
 5 ATGCCTCGG GTTAAATTGA ATAGAAATCG GACTTATGAA CGAAATATGT TTAAGTCGAA 2220  
 CTCCTTGTTT ATACTTATAA ATTITACGGG TTAAATATAA TACTTATTTA CCTGTAATAT 2280  
 ATGCATAATT nCTTCAGTCG GTCAGCCTGT CGTTGCATAG TTCTATGCA GCAATATGCAT 2340  
 10 ATCCTAATCC TTAAACATTG GCATTnCTGC AAATGAACGC ATAGAATCCA TTCATGTTA 2400  
 ACTTTTThCA ACAATATGTCT nACATG 2426

(2) INFORMATION FOR SEQ ID NO: 543:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1874 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 543:

GAGTTGGGA ATGTGCTCAA AATATGCGGA CTTTATGCAT TyCGGAATTG sCCaATTGCA 60  
 25 GCTTTAAGCT ATGGTCAAAA AAAGAGGGTC ACTATAGCAT CTGTTTATG CTTAAATCCG 120  
 GAAATAATCA TATTGGATGA ACCGACTGCT GGTCAAGATT TCTATCATT TAATGAGATA 180  
 ATGTCAATTT TAATTGAACT AAACAGACAG GGAAGACGA TTATTATGAT TACGCATGAT 240  
 30 ATGCATTTAT TGTCAGASTA TAGITCAAGA ACAGTTGTAT TATCAAAAGG TCAAGTCGTT 300  
 GCTGATACCA CGCCAgTATT GGTTTTAAAT GATAAAAAA TCTGTGAGAT TGCAATCATTG 360  
 AGACAAACAT CGCTATTGTA AATGGCCGAA TATATAGGGA TTAGCGAGCC ACAGAAATTA 420  
 35 GTACAATTAT TTATTAACCA TGATAGGAAG GTGAGAGGcC AATGAATCAA TATAATACTA 480  
 TAGGTTTICA CCCGGGAAAT AGTCGTATTC ATCAATTAAA TGCGACTGTT AAACITTTAT 540  
 40 TCTTATTAGT TGTTTCTATT TCTGCAATGG TGACTTATGA CACAAGATAT TTAATTTTAA 600  
 TTATGTGCTC ATCTATTTTA TTGGTCAAAAT ATGCTCATAT TGAATGGAAA CAAGTTCGCT 660  
 TTGTTGTAA ATTCAATCTG TTTTTCACAA TAATAAATAT TAITGCCGTG TACATATTTG 720  
 45 ACCCTGAATA TGGTGGAAG ATTTATAATC AGCGTACAGA GTTAGTCAAT GGTATTTGTC 780  
 GATTACGCT AACATCACAG GAATTATCT ATCTTTTTAA TCTAATATTA AAATATATTA 840  
 GTACAGTTCC TTAGCGTTA ATATTTTTAT TCACAACGAA TCCGAGTCAT TTTGCTGCAA 900  
 50 GTTTAAATCA GCTAGGTGTG AATTATAAAA TCAGTTACGC AGTCTCATA GCATTAAAGT 960  
 ATATTCCAGA TATTCAAGAA ACATATTtTA ATATTTCACA AGCGCAACAA GCAAGAGGAT 1020

# EP 0 786 519 A2

TACCTTAAAT ATTTTCTAGT ATCGAAAGAA TTGACACTAT TAGTACTGCT ATGGAGTTAA 1140  
 GACAATTCGG ACAGTATAAA AGGAGAACCT GGTACGTCAA AAAACAATTA AAAAAAGATG 1200  
 5 ATTATGTTGT TTTGTGTTTG ACGTTAATAC TTCTGATGTT AGTAGTTACA TTATTCTTTT 1260  
 TAAATAATAG TCGATATTTT AACCCGTGGC ATTAGTATTC ATATAAATAG TCTTTAAATA 1320  
 GAAATAGGAG GGAGACATTT AATGATAAAT ACTGAAAGAT TAAATTTAAT GATTCCAAGT 1380  
 10 TCCTCGCATT TAAATTGAACT TTATAATATT TGTAGTCATC CACAAGCAA TATATACACT 1440  
 CCCAAAGGTT TACATAATTC CAAATTAGAC ACACAACGGT GGATTGAAAA ATGGCGAAAC 1500  
 CATTGGCAAC AATATCAATT TGGTTACTTT GTATTGGTAA AAAAAATAGA TTGTAGTGTT 1560  
 15 ATTGGTATTT GTGGATATGA ATATCGACAA TTAAAGCAAG AaACAGTATT AAATTTATTT 1620  
 TATAAATTAC ATCCAAGTTT TGAAGGACAA GGGTACGCAT GTGAGGCTAT TACAGCAATC 1680  
 20 ACAAATTTTG TGAATTATAT CGATCAAGAA ACAGTAAAAG TTATCAGGAC AAATAAGTGT 1740  
 AACCAACGTT CAATAAATTT AGCAGAAAGG CTTAAATTCa AGCGAGACGA TACTATGGAC 1800  
 GACATTATCA ATCAAGGAGA TATTGTGTTT TAAaAATAAA ATACTATGAC ATTATCTAAA 1860  
 25 AAATAAAATT AAAA 1874

(2) INFORMATION FOR SEQ ID NO: 544:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 5280 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 544:

TCAACATTC TAACACCAAT GTGAAAATGA TCTATGTGAT TTGCAATGGC TTGATTGTGA 60  
 40 ATATGTGTGC CTAATGACC GTAGCACCT GTTAACATAA TATTCATTCA CTCATCTCC 120  
 TAATCTTTAT ATACATAACA TAATACTTAT TTGATGGTTT TCAAAACATT TGATTTTATA 180  
 45 AAAAAATCTA ATCTGTATTT ATTGTGACG TGTATAGTAA ATACGTAAAT ATTATTAATG 240  
 TTGAAAATGC CGTAATGACG CGTTTATGTT GATGTGTATC ACTAATATCA TTGAAAATTT 300  
 TAATCaGGTA CTACGACAAT ATGATGTCTG TTTTGTGTCT GAAAGTTTTA CAGTTTTTAA 360  
 50 AATAAAAAAT GTATAAAGTG TGATTGTGAT AAAAAAGAGT CTCACGGAT AAGAATTGAT 420  
 TAATAACAGT TAGCATTTTA TTAATTACCT TAACAATGAT TCAAGTTTAG TTAATGAGG 480  
 TTTAATTTGA AAGGGGATAG CGCCTCAATA TAATGTAGGT AGATTGTTCA TATTACGTAA 540

55

|    |             |            |            |             |             |            |      |
|----|-------------|------------|------------|-------------|-------------|------------|------|
|    | AAGCAAGTTT  | AGCATTGGGA | ATGTTAGCAA | CAGGTGTAAT  | TACaTCGAAT  | GTACAATCAG | 660  |
|    | TACAAGCGAA  | aGCAGrAGTT | AaaCAACAAa | GTGAaTCAGA  | GTTAAACAC   | TATTATAATA | 720  |
| 5  | AmCCAATTTT  | AGAGCGTAAA | AATGTGACTG | GATTTAAATA  | TACTGTAGAG  | GGTAAACACT | 780  |
|    | ATTTTAGAAGT | CACAGTAGGG | CAACAGCATT | CTCGAATCAC  | TTTACTTTGGA | TCTGATAAAG | 840  |
| 10 | ATAAATTATA  | AGACGGAGAA | AACTCAAATA | TAGATGTGTT  | TATCCTTAGA  | GAAGGTGACA | 900  |
|    | GTAGACAAGC  | AACAAATTAC | TCAATTGGTG | CGGTTACAAA  | ATCAAAATAGT | GTGCAGTATA | 960  |
|    | TTGATTATAT  | CAATACGCCA | ATTTTAGAAA | TCAAGAAAGA  | TAATGAAGAT  | GTACTTAAAG | 1020 |
| 15 | ATTTTTACTA  | CATTTCAAAA | GAAGACATCT | CATTAAAAGA  | ACTTGATTAT  | AGATTAAGAG | 1080 |
|    | AACGTGCGAT  | TAAACAACAC | GGCTTGATT  | CAAAATGGTCT | TAAACAAGGT  | CAAATTACAA | 1140 |
|    | TTACAATGAA  | TGATGGCACA | ACACATACAA | TCGATTTAAG  | TCAAAAACCT  | GAAAAAGAAC | 1200 |
| 20 | GTATGGGTGA  | GTCAATCGAC | GGCACTAAGA | TTAATAAAAT  | TCTAGTAGAA  | ATGAAATAAT | 1260 |
|    | ACTTCTTAAC  | AACAAGCGC  | TATGTTGAAT | AGTGCTTGTT  | ATGGAAATAT  | ATGGAAGTTA | 1320 |
|    | AGCGACGTAC  | TGTTGCTTAG | CTTCTTTTTT | TGAGGGGAAA  | AGTTACAAAA  | CTCACACAAA | 1380 |
| 25 | CAGTCGCACC  | ACGCATTATC | TTTTGCTTAA | ATAGCTTAAT  | CATATTTTAT  | GAATAGTTAA | 1440 |
|    | AAACAGGTTA  | ATGTGAATAT | CCGAATACAG | CTCCTATAAT  | ATGGGTGTAT  | GATTCAAATT | 1500 |
| 30 | ACGTATAAAA  | ACAATCTAAT | TATAATAGAT | TGGAGCATAC  | AACTATGAAA  | ATGAAAAATA | 1560 |
|    | TTGCAAAAAA  | AAGTTTGTTA | TTAGGAATAT | TAGCAACAGG  | GTAAACACT   | ACAACGGAAA | 1620 |
|    | AACCAGTTCA  | TGCCGAAAAA | AAACCTATTG | TAATAAGTGa  | AAATAGCAAA  | AAATTAAGAG | 1680 |
| 35 | CTTATTATAA  | TCAACCTAGT | ATTGAATATA | AAATGTGAC   | AGGTTATATC  | AGTTTCATTG | 1740 |
|    | AACCAAGTAT  | TAAATTTATG | AATATCATAG | ATGGTAATTC  | TGTTAATAAT  | ATTGCTTTAA | 1800 |
|    | TTGGCAAAGA  | TAAGCAACAT | TATCATACGG | GTGTACATCG  | TAATCTTAAT  | ATATTTTACG | 1860 |
| 40 | TTAATGAGGA  | TAAGAGATT  | GAAGGTGCAA | AGTACTCTAT  | TGGGGGTATC  | ACGAGTGCAA | 1920 |
|    | ACGrTaAAGC  | TGTCGACCTA | ATAGCAGAAG | CAAGAGTTAT  | TAAAGAAGAT  | CATACTGGTG | 1980 |
|    | AATATGATTA  | TGACTTTTTT | CCATTTAAAA | TAGATAAAGA  | AGCGATGTCA  | TTGAAAGAGA | 2040 |
| 45 | TTGATTTTAA  | ATTAAGAAAA | TACCTTATTG | ATAATTATGG  | TCTTACCGT   | GAAATGAGTA | 2100 |
|    | CAGGAAAAAT  | TACAGTCAAA | AAGAAATACT | ATGGAAAGTA  | TACATTTGAA  | TTGGATAAAA | 2160 |
| 50 | AGTTACAAGA  | AGACCGTATG | TCCGATGTTA | TCAATGTCAC  | AGATATTGAT  | AGAATTGAAA | 2220 |
|    | TCAAAGTTAT  | AAAAGCATAA | CACATATACT | TGATGACGAA  | ATAAGTTGAA  | ATTGAAATAG | 2280 |
|    | AGAGGTTAAG  | TGACGATCAA | ACGTTGCTTA | ACTTCTTTTT  | AATGCTTAAA  | AATTATTCTA | 2340 |

55

|    |             |             |             |             |             |              |      |
|----|-------------|-------------|-------------|-------------|-------------|--------------|------|
|    | TTAATAATAC  | TTCAATAATT  | GTTAAAGGGG  | GTTTAATGTG  | ATTATCTTAG  | AACGCCATCT   | 2460 |
|    | ATAATGATGT  | TGTAATGATTC | AAATTACGTA  | AAAAAGCAAT  | CGAATATAAT  | ATAGATTGGA   | 2520 |
| 5  | GCATACAAAT  | ATTGAAATGA  | GAACAAATGC  | TAAAAACAGT  | TTAGCACTAG  | GGCTTTTAAAC  | 2580 |
|    | AACAGGCGCA  | ATTACAGTAA  | CGACCGCAATC | GGTCAAGACA  | GA AAAAATAC | AATCAACTAA   | 2640 |
| 10 | AGTTGACAAA  | GTACCAACGC  | TTAAAGCAGA  | GCGaTTAGCA  | ATGATAAAAC  | TAACAGCAGG   | 2700 |
|    | TGCAAAATTC  | CGCAACAAC   | AAGCAGCTAA  | CACAAGACAA  | GAACGCACGC  | CTAAACTCGA   | 2760 |
|    | AAAGGCACCA  | AATACTAATG  | AGGAAAAAAC  | CTCAGCTTCC  | AAATATGAAA  | AAATATCACA   | 2820 |
| 15 | ACCTAAACAA  | GAAGAGCAGA  | AAACGCTTAA  | TATATCAGCA  | ACGCCAGCGC  | CTAAACAAGA   | 2880 |
|    | ACAATCACAA  | ACGACAACCG  | AATCCACAAC  | GCCGAAAACT  | AAAGTGACAA  | CACCTCCATC   | 2940 |
|    | AACAAACACG  | CCACAACCAA  | TGCAATCTAC  | TAAATCAGAC  | ACACCACAAT  | CTCCAACCAT   | 3000 |
| 20 | AAAACAAGCA  | CAACACATA   | TGACTCTTAA  | ATATGAAGAT  | TTAAGAGCGT  | ATTATACAAA   | 3060 |
|    | ACCGAGTTT   | GAATTTGAAA  | AGCAGTTTGG  | ATTTATGCTC  | AAACCATGGA  | CGACGGTTAG   | 3120 |
|    | GTTTATGAAT  | GTTATTCCAA  | ATAGGTTTCAT | CTATAATGCT  | GCTTTAGTGT  | GA AAAAGATGA | 3180 |
| 25 | GA AAAAATAT | AAAGATGGAC  | CTTAGCATAA  | TATCGATGTA  | TTTATCGTTT  | TAGAAGACAA   | 3240 |
|    | TAAATATCAA  | TTGAAAAAAT  | ATTCTGTCGG  | TGGCATCAGC  | AGACATAATA  | GTAAAAAAGT   | 3300 |
| 30 | TAATCACAAA  | GTAGAATTAA  | GCATTACTAA  | AAAAGATAAT  | CAAGGTATGA  | TTTCACGCGA   | 3360 |
|    | TGTTTCAGAA  | TACATGATTA  | CTAAGGAAGA  | GATTTCTCTG  | AAAGAGCTTG  | ATTTTAAAT    | 3420 |
|    | GAGAAAAACA  | CTTATTGAAA  | AACATAATCT  | TTACGGTAAC  | ATGGGTTTAC  | GAACAACTCG   | 3480 |
| 35 | TATTAAAAAT  | AAAAACGGTG  | GGAAATATAC  | GTTTGAATTA  | CACAAAAAAC  | TGCAAGAGCA   | 3540 |
|    | TCGTATGGCA  | GACGTATAG   | ATGGCACTAA  | TATTGATAAC  | ATTGAAGTGA  | ATATAAAAAA   | 3600 |
|    | ATCATGACAT  | TCTCTAAATA  | GAAGCTGTCA  | TCGGA AAAAC | AAGAAGTTAA  | GTGCAACCGG   | 3660 |
| 40 | TTTACATGTT  | GCTTAGCTTC  | TTTTATTATG  | CGTAATGATG  | TAAAAAGACG  | AATATTTCATT  | 3720 |
|    | TGTTTGTAATA | AGTGGCATTT  | CTATGCTCTA  | AAAGTGACGA  | AACTTCAAAT  | GTGCCAAGTG   | 3780 |
|    | TTGAATCACA  | TCAAAATCAT  | TTTTATTATTA | CGAACATTAT  | GGATTCTCTA  | ATTTACTTTAA  | 3840 |
| 45 | CGATGATTCA  | AATATAGTTA  | AACAAGGTTT  | AATGTGAATG  | GAGCAATACG  | CCATCTATAA   | 3900 |
|    | TAAAGCTGTA  | TGATTCAATG  | AATGTAAATCG | AACAAATCTA  | ATAATTACGA  | ATGGAGCATA   | 3960 |
| 50 | CAACTATGAA  | AATAACAACG  | ATTGCTAAAA  | CAAGTTTAGC  | ACTAGGCCIT  | TTAAACAACG   | 4020 |
|    | GTGTAATCAC  | AACGACAACG  | CAAGCAGCAA  | ACGCGACAAC  | ACTATCTTCC  | ACTAAAGTGG   | 4080 |
|    | AAGCACCAACA | ATCAACACCG  | CCCTCAACTA  | AAATAGAAGC  | ACCGCAATCA  | AAACCAAAACG  | 4140 |

EP 0 786 519 A2

|    |  |      |
|----|--|------|
|    | CGCCTTCAAC TAAAGTGACA ACACCTCCAT CAACAAACAC GCCACAACCA ATGCAATCTA  | 4260 |
|    | CTAAATCAGA CACACCACAA TCGCCAACCA CAAACAAGT ACCAACAGAA ATAAATCCTA   | 4320 |
| 5  | AATTAAAGA TTTAAGAGCG TATTATACGA AACCAAGTTT AGAATTTAAA AATGAGATTG   | 4380 |
|    | GTATTATTTT AAAAAATGG ACGACAATAA GATTATGAA TGTGTGCCA GATTATTTCA     | 4440 |
| 10 | TATATAAAAT TGCTTTAGTT GGTAAGATG ATAAAAATA TGGTGAAGGA GTACATAGGA    | 4500 |
|    | ATGTCGATGT ATTTGTCGTT TTAGAAGAAA ATAATTACAA TCTGGAAAAA TATTCTGTCTG | 4560 |
|    | GTGGTATCAC AAAGAGTAAT AGTAAAAAAG TTGATCACA AGCAGGAGTA AGAATTACTA   | 4620 |
| 15 | AGGAAGATAA TAAAGGTACA ATCTCTCATG ATGTTTCAGA ATTCAAGATT ACTAAAGAAC  | 4680 |
|    | AGATTTCCCT GAAAGAAGCT GATTTTAAAT TGAGAAAAA ACTTATTGAA AAAAATAATC   | 4740 |
|    | TGTACGGTAA CGTTGGTTCA GGTAATAATTG TTATTAAAAA GAAAAACGGT GGAAAGTACA | 4800 |
| 20 | CGTTTGAATT GCACAAAAA TTACAAGAAA ATCGCATGGC AGATGTCATA GATGGCACTA   | 4860 |
|    | ATATTGATAA CATTGAAGTG AATATAAAAT AATCATGACA TTCTCTAAAT AGAAGCTGTC  | 4920 |
|    | ATCGGAAAAA CAAGAAGTTA AGTGACAACG GCCTACATGT TGCTTAGCTT CTTTGTGTAT  | 4980 |
| 25 | GTTTCGATGAT TTGAGAAGCC GAATTTTCGA TGGGTCCAAA TATGACGTGG AAGAGACCTG | 5040 |
|    | AATTATCTCT TAAATCCCTA TCTATCGGGT GTGAAGCACA ACGGATCAG TTTTATTAA    | 5100 |
| 30 | CGAACATTAT AGATTCCTTA ATTTACTTAA TAATGATTCA ATGATTATTA AACATGGTTT  | 5160 |
|    | AATGTGAAAG GTCAATACG CTAACATATA TAAAGCTGTA TGATTCAATA GACATGAGCG   | 5220 |
|    | AACAAATCTA ATAATTACGA ATGGAGCATA CAACTATGAA AATGACAGCA ATTGCGAAAG  | 5280 |

35 (2) INFORMATION FOR SEQ ID NO: 545:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 886 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 40 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 545:

|    |  |     |
|----|--|-----|
| 45 | AGTAAATTTG CCGGTATGAT GGACACAAAC GGTGACCTTG GTCAAGGTGA ATTACGATT   | 60  |
|    | AATCCACCTA AATCAGATTT gAACGAATTA CCTTGGGCTA CACGTAAAAA TAAACAGCCA  | 120 |
| 50 | GCTTCATCCG AAAAAgGTTc AAGTGGTCAT CATGGTAATG CAGCAATGCC TCAAAACCa   | 180 |
|    | TTAGATTATC AAATATCTAT TGATAAGGTC GTTGAACAGG CGCAAAAAGC TGGTATTAAA  | 240 |
|    | AAGCCGTTTT CAATCGTATA TCCAAGTGAT AAAAAATGGTA CCTTTATTGT ATCTAATACT | 300 |

55

GATCAATATA GCGGTAAAAA GCTAGGTACG ATTAANTATG ATGACTACGG TATTATTGCT 420  
 AAATGSGTTA CATGGGGCAT TCCGCTTCAC GAAGGTCATT TATTCGGCAT TTTAAATAAA 480  
 5 ATCATTAAIT TATTGTGATG TATCGCTTTA TTAGTAGCCA TTGGCATGGG GTTTGTCTCT 540  
 TGGATAAAGC GTACAAAAAA TACTGCAGTA AAAGTACCAC ATCGCGTAAA AAAACGAGCA 600  
 TCTATATCAC TCATAATATG TTTAATTGTA TTAGGATTAT TAATGCCATT ATTTGGATTA 660  
 10 TCACITATCC TTGTATTAT AATTGAATTA ATATTATATA TTAAAGATCG TCGTGCTAAA 720  
 CAATAATGCA CTTAAAGTTT TGAAGTACG AATTTTACAA AATGgATTCT CGTCTCTCTA 780  
 ATTACTTAAA ACGGGgTtCy AaTAATAAAT CgTACTGaTG GgAAAGTTTT TACTTTTTAT 840  
 15 CTGtCCGAtT TTTTnGAaWt TGAAGATAAA AAAGCATCTA AAACGC 886

(2) INFORMATION FOR SEQ ID NO: 546:

20 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 4336 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 546:

GGCATTGTG TCCTTATATA AGGAACGTG tTAAATACAT TACTGTTGTT AAGTTGTTTT 60  
 30 TGTAATTCAA AGAGCAGAAC AGAGTAACAT CATCAGTTGT AGTAAACGAT AATCCGGTAA 120  
 AACAACATAA TGAATAATG AAAGTCATT AACCTGAACA TTAATAATATA TTGTTTTTTC 180  
 35 ATTAAGAATA ATTCAAGTAT ATTTAAATCG AGGTTAATTA TCGTATGAAA CGATGCACGT 240  
 TATAATAAAA ATGTATGATT CAAATTACGT AATGAAAACA ATCCAATATA TTAAGATTGG 300  
 AGCAATAAAA TATGAAATTT ACAGCATTAG CAAAAGCGAC ATTAGCTTTA GGAATTTTAA 360  
 40 CAACAGGAAC TTTAACAACA GAAGTTCATT CAGGTCATGC AAAACAAAAT CAAAAGTCAG 420  
 TAAATAAACA TGACAAGGAA GCATTATACC GATACTACAC TGAAGAGACT ATGGAATGA 480  
 AAAATATTAG TGCTTTGAAA CATGGTAAAA ACAACTTACG TTTTAAGTTT AGAGGTATTA 540  
 45 AGATTCAAGT TTTACTGCCT GGAATGATA AAAGTAAAT TCAACACGCT AGTTATGAGG 600  
 GGTTAGATGT TTTCTTTGTT CAAGAAAAAA GAGATAAGCA CGATATATT TATACTGTTG 660  
 50 GTGGTGAAT ACAGAAATAA AAAACATCTG GAGTTGTCAG TGCACCAATA TTAATATTT 720  
 CAAAAGAAAA GGGTGAAGAT GCTTTTGTGA AAGGTTACCC TTATTACATT AAAAAGAAA 780  
 AAATAACACT AAAAGAAGT GATTATAAGT TGAGAAAGCA TCTAATTGAA AAATACGGAC 840

|    |  |      |
|----|--|------|
|    | ATAACCTTGA TTTAAGATCT AAATTAAAA TTAATATAT GGGGAAGTC ATAGAAAGCA     | 960  |
|    | AACAAATTA AGATATTGAA GTTAACCTAA AGTAATCAT TACGAATAAT TAAAGTAAT     | 1020 |
| 5  | TGAAGCGGCT TAACGGTGAA ATGTAAATTG GTGGCATAG CTTATACAAA AAGGATGCAT   | 1080 |
|    | CAATCGATAT CGTCGTTAAG CCGTTTGGT TTGTGTGCA TGAATCCTAT CCCAATCTCC    | 1140 |
| 10 | ATAAAGGTAA AATTTCACAC ACCAACATCA AAATCTCCA CATCGCAACA TAACCAAATG   | 1200 |
|    | TTATAATAAA TCTATTACAC AAAGAGATAA ATTACTTATT CAAAGCGGA GGAATCACAT   | 1260 |
|    | GTCTATTACT GAAAAACAAC GTCAGCAACA AGCTGAATTA CATAAAAAT TATGGTCGAT   | 1320 |
| 15 | TGCGAATGAT TTAAGAGGGA ATATGGATGC GAGTGAATTC CGTAATTACA TTTTAGGCTT  | 1380 |
|    | GATTTTCTAT CGCTTCTTAT CTGAAAAAGC GGAACAAGAA TATGCAGATG CCTTGTGAG   | 1440 |
|    | TGAAGACATC ACGTATCAAG AAGCATGGGC AGACGAAGAA TACCGTGAAG ACTTAAAAGC  | 1500 |
| 20 | AGAATTAATT GACCAAGTCG GTTACTTCAT TGAGCCAGAA GATTATTCA GTGCGATGAT   | 1560 |
|    | TCGTGAAATT GAAACGCAAG ATTTGATAT CGAACACCTG GCGACGGCAA TTCGTAAGT    | 1620 |
|    | TGAAACATCA ACATTAGGTG AAGAAAGTGA AAATGACTTT ATCGTCTGT TCACGATAT    | 1680 |
| 25 | GGATTTGAGT TCAACGCGAC TAGGTAACAA TGTCAAAGAA CGTACTGCTT TAATCTCTAA  | 1740 |
|    | AGTCATGGTT AATCTTGACG ACTTACCATT CGTTCACAGT GACATGGAAT TTGATATGTT  | 1800 |
| 30 | AGGTGATGCA TATGAATTC TAATTGGGCG CTTTGGGCG ACAGCGGTA AAAAGCAGG      | 1860 |
|    | CGAGTCTAT ACACCACAAC AAGTATCTAA GATACTGGC AAGATTGTCA CAGACGGTAA    | 1920 |
|    | AGATAAATTA CGTCACGTGT ATGACCCAAC ATGTGGTTCA GGTTCACGTG TGTTACGTGT  | 1980 |
| 35 | TGTTAAAGAA ACACAAGTGT ATCGTTATT CGGTCAAGAA CGTAACAATA CTACATACAA   | 2040 |
|    | CTTAGCACGC ATGAATATGT TATTACATGA TGTGCGTTAT GAGAACTTCG ATATCCGTAA  | 2100 |
| 40 | TGATGACACA TTGGAAGAAC CAGCCTTTT AGGCAATACA TTTGATCGCG TTATTGCGAA   | 2160 |
|    | CCCACCGTAT AGTGCAGAAAT GGAATGCAGA TTCAAAGTTT GAAATGACG AACGATTCAG  | 2220 |
|    | TGGTTACGGC AAACCTTGCGC CTAAGTCTAA AGCAGACTTT GCCTTTATTC AACACATGGT | 2280 |
| 45 | ACATTACCTA GACGATGAAG GTACCATGSC CGTTGTACTC CCACATGGTG TATTATTCG   | 2340 |
|    | AGGTGCTGCA GAAGGTGTCA TTCGTGTTA TTAAATTGAA GAAAGAAGT ACTTGAAGC     | 2400 |
|    | TGTGATTGGT TTGCCAGCGA ATATTTCTA TGGGACAAGT ATTCCAACAT GTATTTTAGT   | 2460 |
| 50 | ATTTAAAAAA TGTGCGCAAC AAGACGACAA CGTACTATT ATCGATGCAT CCAATGATTT   | 2520 |
|    | TGAAAAAGGA AAAAAACAA ATCATTTAAG CGATGCCCAA GTCGAACGTA TTATAGACAC   | 2580 |
| 55 | ATATAAGCGT AAGGAAACAA TTGATAAATA TAGCTACAGC GCGACACTAC AAGAGATTGC  | 2640 |

EP 0 786 519 A2

|    |   |      |
|----|---|------|
|    | GATTGATTTA GATCAAGTCC AACAAAGATT GAAAAATATC GATAAAGAAA TCGCAGAAAT   | 2760 |
|    | TGAGCAAGAA ATCAATGCAT ACCTGAAAGA ACTTGGGGTG TTGAAAGATG AGTAATACAC   | 2820 |
| 5  | AAAAGAAAAA TGTGCCAGAA TTGAGGTTCC CAGGGTTTGA AGCGAATGG GAAGAGAAGC    | 2880 |
|    | AGTTAGGGGA TCTTACAGAT AGAGTAATTA GGAATAATAA AAACCTTAGAA TCGAAAAAGC  | 2940 |
| 10 | CTTTAACAAAT ATCCGGACAG TTAGGTTTAA TTGATCAAAAC AGAATATTTT AGTAAATCAG | 3000 |
|    | TTTCGTCGAA AAATCTAGAA AATTATACAC TAATAAGAAA TGGAGAATTC GCGTATAACA   | 3060 |
|    | AAAGTTATTTC TAATGGATAC CCATTAGGGG CTATTAAAAA ATTAACTAGA TATGATAGTG  | 3120 |
| 15 | GTGATTGTCT CTCCTTGTAT ATTGTGTTTT CTATTAAAAA TGAAATGTCT AAAGACTTCA   | 3180 |
|    | TGGAAGCATA TTTTGATTGC ACACACTGGT ATAGAGAAGT TTCTGGAATT GCAGTTGAGG   | 3240 |
|    | GTGCAAGAAA TCACGGATTA TTAATGTTT CTGTGAATGA TTTTTTACT ATTCTAATTA     | 3300 |
| 20 | AATATCCAAAG TTTAGAAGAA CAGCAAAAAA TAGGCAAGTT CTTACAGCAA CTCGACCGAC  | 3360 |
|    | AAATTGAATT AGAAGAACAA AAGCTTGAAT TACTTCAACA ACAGAAAAAA GGCTATATGC   | 3420 |
|    | AGAAAAATTT CTCACAGGAA CTGCGATTCA AAGATGAGAA TGGTGAAGAT TATCCAGATT   | 3480 |
| 25 | GGGAAAAATAG CAAAATAGAA AAATATTTAA AAGAGAGAAA CGAACGTTCT GACAAAGGGC  | 3540 |
|    | AAATGCTTTC AGTAACTATA AATAGTGGCA TTATAAAATT TAGTGAATTG GATAGAAAAG   | 3600 |
| 30 | ATAATTCAAG TAAAGATATA AGTAATTATA AAGTAGTTAG GAAAAATGAT ATTGCATATA   | 3660 |
|    | ATTCTATGAG AATGTGGCAA GGGGCTAGTG GTAAATCAAA TTATAATGGG ATTGTTAGCC   | 3720 |
|    | CTGCATATAC TGTGCTTTAT CCAACACAAA ATACTAGCTC ATTATTTATT GGATATAAGT   | 3780 |
| 35 | TTAAAAACACA TAGAATGATT CATAAATTTA AAATTAAATC ACAAGGATTA ACATCAGATA  | 3840 |
|    | CATGGAACCT AAAATATAAA CAATTAAAAA ATATAAATAT AGATATACCT GTATTGGAGG   | 3900 |
|    | AACAAGAAAA GATAGGTGAT TTCTTTAAAA AAATGGATAT ATTGATAAGT AAACAGAAAA   | 3960 |
| 40 | TGAAAATTGA AATATTAGAA AAAGAGAAAC AATCCTTTTT ACAAAAAATG TTCTTATAAC   | 4020 |
|    | TTTGATAAAT ACATAGATTG CATAAGAATA AAATTGTGAT AATTTAACAT AAAAGTTGTA   | 4080 |
|    | AAAGTAAAGT GAATTAAAAA CGAACATTAA ATTTAGGCAC TGTGAAAGCG CAGTGTCTTT   | 4140 |
| 45 | TTTGTGTCGA AATTGTGTAC AGAATAAGTA GTTAAATAAA GATTAAAGTG AGATAAAGTG   | 4200 |
|    | TTATTCGTAA ATAAAAGAGA GTAGATCGAT AGGAATTGAA TGATATTAGT TAACTATTTA   | 4260 |
| 50 | TTAAATTAAT TAATAATGAT TAATTTTITAG TTAAAGTAAG TTTAATGTGA AGCACGACCA  | 4320 |
|    | TTGCTCATTAA TAATGA  | 4336 |

(2) INFORMATION FOR SEQ ID NO: 547:

55



# EP 0 786 519 A2

(A) LENGTH: 487 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 547:

|    |   |     |
|----|---|-----|
| 10 | TAAGCTATCT GAGATAATTG CTGATAACAT TAAACCGGCA ATTCAGGTT TAATTTCAAA  | 60  |
|    | GCCACGTTCT CTAACATTT TGTATAAAT TGTAGCTGTA CAACCACTG GTTCAGCAG     | 120 |
|    | ATAACATAAA GGACCAGCAG TTTCGAAATT TGCAATTCTG TGATGATCAA TTACATGCTT | 180 |
| 15 | AATTGTAGCA GAGGCAATCG TATCAGAACT TTGTTGGaAT TCGTTATGAT CAACTAAGAT | 240 |
|    | AACATCTTGA CCATCTAAAT CATCTGTTAA TAATTCGGGA GCAGGTACAT TAAATGTATC | 300 |
|    | TAACGCGAAT TGAGTTTCTG CACTCACATC ACCTAAACGG TATGCTTTGG CTCCTGAATT | 360 |
| 20 | ACCTCGAAGT TGTTCaAATT CTGCCaTAAT AATCGCAGAT GAAATTGCAT CAGTGkCTGG | 420 |
|    | ATTCTTATGT CCGAAAATAT ATGTTTTAGC CAnTGTCAAA TATCTCCCTT GTAAATTGTA | 480 |
| 25 | TTCTTTA   | 487 |

(2) INFORMATION FOR SEQ ID NO: 548:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 871 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 548:

|    |  |     |
|----|--|-----|
| 35 | TTGGTGGTGC AGCAGTTATA GCAATTGTTT TAGCATTATC TGGTAAGTTC ACTGCATTAA  | 60  |
|    | TTTCTTCTAT ACCTACmCCA GTTATGGGAG GAGTATCTAT ATTACTTTTC GGTATTATTG  | 120 |
| 40 | CAGCAAGTGG CTTAAGAATG TTAGTTGAAA GCAAAGTAGA TTTTGCGAAC AATCGAAATT  | 180 |
|    | TAGTTATAGC TTCTGTAATT TTAGTTGTAG GTATCGGTAA TTAGTATATT AACTTAAAG   | 240 |
| 45 | AAATTGGTAT CAACCTTCAA ATTGAGGGGA TGGCATTAGC TGCACTTTCA GGAATTATTT  | 300 |
|    | TGAACCTTAAT CTTACCTAAA GAGAAAAAAC AAAACAATTA AGATTTACAA ATTAAGGAGG | 360 |
|    | GGCCTTTTAT GAATCATTTA TTATCAATGG AACATTATC TACAGATCAA ATATACAAAC   | 420 |
| 50 | TTATCCAAAA GGCAAGTCAA TTTAAATCTG GTGAACGTCA ACTACCAAAC TTTGAAGGGA  | 480 |
|    | AaTATGTCGC AAATTTTATC TTGAAAAATT CTACTGfAC AAAATGTAGT TTTGAAATGG   | 540 |
| 55 | CAGAACTTAA GCTAGGGTTA AAAACGATTA GCTTTGAAAC ATCAACATCA TCTGTTTCAA  | 600 |

EP 0 786 519 A2

TCATTAGACA TCCGTTTAAAT AACTACTATG AAAAATTAGC GAATATTAAC ATCCCAATTG 720  
CGAATGCTGG TGATGGTAGT GGACAACATC CAACACAAAG TTTACTTGAT TTAATGACGA 780  
5 TATATGAAGA ATATGGATAT TTTGAAGGCT TGAATGTATT GATTGTGGA GacATTaAAA 840  
ATTCACGTGT CGCACTAGTA ATTACCAAG T 871

(2) INFORMATION FOR SEQ ID NO: 549:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 400 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
15 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 549:

20 TCAATCTAAT ATATACTTCA TGACTTCCCG CCAAAGGCGC AATCGTAGGA TTAGTCTGTT 60  
TATCGATAAG ATCTATTAAAT ATTGTTTCTA CTTTAGATTG ACCTATTCCC GCAAAATCTTA 120  
ATAGTTCAGA ATGTATAATT CGATTATGGT TTATAAAATG TGACAACAAT TCATTTTTCa 180  
25 CCATTGGTTG CATTTCCTTC GGTGGACCTG GTAATAAAAT AATTGTGTTG TTTTCAAAAT 240  
TCACCATCAT TCCTGGAGCC ATGCCATGAT GATTGTGTAA TACAGTTGAA CCTTCAATTA 300  
CTAAAGCCTG TTGTCTATTA TTAGGTGTCA TTTCTGTCC TTGTCCTCA aAAwAGCTTT 360  
30 CAATATATTG AAAGAAGGCT CATCAATAAC TAAATCTnTA 400

(2) INFORMATION FOR SEQ ID NO: 550:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1523 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
35 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 550:

ATACCTCCAG CTAGAATACC AGCGTATTTT ATAAAATACT TCCTCCATTG AACTATATCT 60  
45 ATATTTAAAT ATTTAAATTT CGTTGCATTT TCCAATTGAA AACTCATTTT AAAATCAAAA 120  
CTCTAAATGT CTGTGTATTA CTTAAAATTA TACATATTTT GCTTATATTT TAGCATATTT 180  
TGTTTAAACC TATATTACAT TATATCAGAC GTTTTCATAC ACAATAATA ACATACAAGC 240  
50 AAACATTCG TTTATTATTT ATATCACTTA ACTAATTAAT TTATAATTTT TTATTGTTTT 300  
TAAGTTATCA CTTAAAAATC GTTTGGCaAA TTCGTTGTGA CGCTGTGCCA TCTTCTAATG 360

55

## EP 0 786 519 A2

TTTGTTTTAA TGCATCAATG AGTGCCTGTT GATTTCACAC AATTGGACCT GGCAACTCTT 480  
 TTTTATAATC CATGTAAAAA CCTCTAAGCT CATCGCCATA TTTATCTAAG TCATATGCAT 540  
 5 AGAAAAATTG CGGACGCTTT AATACACCGA AGTCGAACAT GACAGATGAG TAGTCGGTAA 600  
 CTAACGCATC GCTGATTAA GATAAATCCG AAATGTCTTC ATAACTGAA ACCTCTTCA 660  
 10 CAAAATCATC ATGTTTCATCA ATACGTGTCA CACTAAATA ATGCATGCGT AAKAAAAATA 720  
 CATATTCATC ATCCAGCGCT TGACGCAAAG CTTCTATATC AAAGTTAACA TTAAATTGAT 780  
 ATGAACCTTC TCGAATCGCT TCATCGTCAC GCCAAGTTGG CGCGTACATA ATCACTTTTT 840  
 15 TATCTAATGG AATATTTAAT CTGTCTTAA TACCATTAA ATATTCAGTA TCATTGCGTT 900  
 TATGTGATAA TTTATCATTT CTGGGATAAC CTGTTTCCAA AATCTTATCT CGACTAACAT 960  
 GAAATGCATT TTGAAATATC GATGTCGAAT ATGGATTAGG TGACACTAGA TAATCCACCC 1020  
 20 GTTGGCTTC TTTTTAAAG CCATCTGGT AATTTTGAGT ATTTGTTCT AGCATTTTAA 1080  
 CGTTACTAAT ATCCAAACCA ATCTTTTTTA ATGGCGTGCC ATGCCATGTT TGTAAGTACG 1140  
 TCGTTGCGGG TGATTTATAT AACCAATCTG GTGTACGTGT GTTAATCATC CwCGCTTTGG 1200  
 25 CTCTGGCAT CGCTAAAAAC CAITTCATTG AAAACTTTGT AACATATGGT ACATTGTGCT 1260  
 GTTGGAAATAT GTGTTATAT CTTTTTTTCA CACCCATAT TAATTGGGCA TCGCATATGTT 1320  
 30 CAGTTAAGTA TTCATATAAT GCTTTGGGGT TGTCGCTGTA TTGTTTACCA TGAAGGCTTT 1380  
 CAAAATAAAT TAGATTCTTG TTGGCAATT TTGATAGTAA TTTAAAGTC GTATATATAC 1440  
 TATGTTCTAT CAATTTTTTA ATTGATTTT TAATCATGTC GTACCTCGA CGTGTTTTTG 1500  
 35 TAATTATATT AATATGATG AGC 1523

(2) INFORMATION FOR SEQ ID NO: 551:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 4923 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 551:

CAGTAAGAGA TTTCTTAAT TGAAAATAAT CTTACTGCTT TTTTAAATTT AATTTCGAGA 60  
 TTCAATATTA GTTTATCTCA TTGTGGCATT AATTGATAAA ATTGTTTTAA TGTATATAAT 120  
 CAAAGTCtCT TCACAATTT TCAATGCTG CATCAGATCT ATGATATGTG AAAAGCTAA 180  
 TTCTTATGCG GTCTAAATGC TCTAAATGGT GTCGATACCT TTCGATGGCT GCAACAATTT 240

EP 0 786 519 A2

|    |  |      |
|----|--|------|
|    | ATGTGGACAA TAAATCTTTT TTCTGCGGAT TATCTAACTC ATAATCAACA TGTGTCACAT  | 360  |
|    | TATAACGTGC TTTTTAGAA AGGCTAGCTA AAAATTGCTC GTGAAAAGCA GTTAATGAAT   | 420  |
| 5  | CTAAATCTAA TTTGATTGTG AATGAGAAAT TGTATTAAAG TAAATATAAG TCGTTTGTAT  | 480  |
|    | AACGCGACAA TTTGTTTAAT ACTTCATAAG CTGTCTAGT CGCTGAACT ACTTCTCTGA    | 540  |
|    | AAAGTATTTT CTTTCTATTC TGTGGTGAA TATGTTTTT TGTAAATGA CGTTCTCCGC     | 600  |
| 10 | TATAGTAATC ATAAATTTTC TCTAACTTTT CGACAGTTG TTTTAAATTA TGACTATCTT   | 660  |
|    | GTTTAATATT ATTAACCTCC GTCGATCAT TTAATACTAA TTTAAACCAC ATAAAAATAT   | 720  |
| 15 | CTGAGGATAT ATTTAATGAA TTATAGTAAA TTTTGTTC AAATTTAGGT GGTAGAAACA    | 780  |
|    | CAAGTTAAC TAGAGATGAA CTTATGACAC CAATCATTAC AAGTACAAAC CTGTAAAAGG   | 840  |
|    | CGGTAATATA GAAAGAACCG GTATGTTGTC CCATAATGAT TAATGCTGTT ACACGCGCCA  | 900  |
| 20 | AAGTAGCAAC ATGTGCTAAA TTAATTTTAA ATAAAAAGC AATAAGTACT ATGACGGTAA   | 960  |
|    | CACCCATAAT GATAAAATTA TCACTAAAAA TTGTTACCAT TGTAACAGAT AGTATGGCAC  | 1020 |
|    | CTATAATGTT ACCCAATGCT TGATCAGAAA CTGTTTTAAA TGAACGATAA ATACTAGGTT  | 1080 |
| 25 | GCATTGCACa ACAGCACTGA CACCAGCTAA GGCTTCAGA CCAACATCAT CCGGTAGTAA   | 1140 |
|    | AGAAGCGATA GACATAGCTA AAATAATGGC TATACCAAGT TTAATAATCC GAGCTCCTAG  | 1200 |
| 30 | TCTCAAAAAA AATGACGCCC CTTTTAAGTT TATTGAATAT CTAATATTCTG TATTCATTAC | 1260 |
|    | TGTTATACAC TTACTAGTTA CAAAATTCAA GCTTATTAT AGTTGTTAAA ATAAATCATA   | 1320 |
|    | CATAATAGTG ATAGCGATGT AAAACTTTAG TCAGAGATTA AAATAGTATA AATTGTGAAA  | 1380 |
| 35 | ATAAAAACTC ACATAGTGAC ATATCAAGTT AAACGTTAAT AGTTAACGAT ATAAATGAA   | 1440 |
|    | TCTACTATGT GAGCAATTGC TTTATTTTAA TTCAAATAAA AATATACTTC CTTAAAAAGT  | 1500 |
|    | ATTTCAATTTG ACTAAAAGCA TAGTCTGCAG CTTTAAAGT TGTGTTAATA TCTTCTCTCG  | 1560 |
| 40 | TATGTTCAGT TGTTAAGAAC CAAGCTTCAA ACTTAGAAGG TGCTAAATTG ATACCTTGAT  | 1620 |
|    | TTAACATTAA TTTGAAAAAT TTACCGAAGC CTTCGCGGTC AGAATGTTCA ACTTGATCAT  | 1680 |
|    | AATGTGTGAC TTTTTCATCT GTAAAGTACA ATGTTAAAGA TCCATAAATA CGATTAAATTG | 1740 |
| 45 | TAGCTGTGAT ATTATGTTTT TCGATTAAAT TAAGTAAACC TTCCTCTAGT TGTGTGCGCTA | 1800 |
|    | AGCTGTCTAA TTTTTCATAA ACACCGTCTT GTTCTAGTAC TTCAGTAAAT GCAATACCTG  | 1860 |
|    | CTTTCATAGA TAACGGGTTA CCAGCCATTG TACCAGCTTG ATATGCAAGT CCTAGAGGTG  | 1920 |
| 50 | CTACTGTGTC CATAATATCT TGACGTCCAC CATAGCCTCC AATGTGTAAG CCACCGCCAA  | 1980 |
|    | CAATTTTACC AAATGCAGTT AAATCAGGGA TAACACCTAA TAAATCTTGA CGCGCACCGT  | 2040 |

55

EP 0 786 519 A2

|    |  |      |
|----|--|------|
|    | AAATTTTCATT AACCTCTTCT AAAAATCCAG GTTGAGGCAT TACCATTCCA AAGTTACCAA | 2160 |
| 5  | CAATTGGTTC TACTAATACT GCGGCAATT CATCACCCCA AAATTCGAAT GCTCTTTAT    | 2220 |
|    | AGCGGTTAAT ATCATTGAAA GGTACAGTAA TGACTTCAGC TCGACGCTT TCTGGAACAC   | 2280 |
|    | CAGCTGAGTC TGGAGAACC AGCTGAGATG GGCCGCTACC TGCTGCAACC AATACTAAAT   | 2340 |
| 10 | CAGAATGGCC ATGATAAGAT CCAGCAAATT TTATAATTTT ATTTCTTTTA GTATATGCAC  | 2400 |
|    | GTGCAACACG AATTGTTGTC ATGACTGCTT CTGTTCAGCA ATTTACAAGG CGAATTTTCT  | 2460 |
|    | CAAGAGATGG AATTGCATCA CGTAATTTT TGCTGAATTC AATTTCTAAT TCAGTCGGTG   | 2520 |
| 15 | TACCAATAA AACACCTTTA GCAGCTTGTT CTTGAATTGC TTTAGTAATA TGAGGATGTG   | 2580 |
|    | CATGCCCCGT AATAATTGGA CCGTATGCTT GAAGGTAATC AATAAATTTA TTGCCATCGA  | 2640 |
|    | CATCATATAA ATATGCACCG TGTCTTCTT TCATAACAAC AGGTGCACCG CCTCCTACAG   | 2700 |
| 20 | CTTTATAAGA ACGAGAAGGG GAATTGACAC CGCTAGAAT ATATTCGTTT GAAAGTGTG    | 2760 |
|    | GTAACGCTTC ACTTTCACCT AAATTCATTT ATATCAACCT CTTTTAATTT AATATTTTCA  | 2820 |
|    | TCTAATATCG TATCATAAAA TTATTATAAT GAAGAAAAAG GTGATTATAT GTTGCAAAAA  | 2880 |
| 25 | GGAGAACAAT TTCCAATATT TAAATTAGAA AATCAAGACG GAACTGTCAAT TACAAATGAT | 2940 |
|    | ACATTAAAAA GTAAAAAGGC GATTATATAT TTTTATCCTA GAGATAATAC ACCTACTTGT  | 3000 |
| 30 | ACCACAGAAG CTGTGACTT TAGAGACAAT TTAGAAATGT TCAATGATTG AGATGTTGCA   | 3060 |
|    | GTATATGGTA TAAGCGGTGA TTCAAAGAAA AAACACCAAA ATTTTATTGA GAAACACGGA  | 3120 |
|    | TTGAATTTTG ATTTATTAGT AGATGAAGAT TTTAAATTAG CTAAGAAAAA GCGGTATATC  | 3180 |
| 35 | AGTTAAAAAA ATCATTTGGC AAAGAAAGTA TGGGCATTGT AAGAACGACT TTTATAATAG  | 3240 |
|    | ATGAACAAGG TAAAGTATTA GATGTTATCG AGAAGGTTAA GGTAAAAACA CAAATAGAAG  | 3300 |
|    | AACTTAAAAA CATTTTGGGG TGACATATAT GAAAGTTGTT GGGTTAAATC GTATCGCTGA  | 3360 |
| 40 | AGTTGAAACT GAATTACAAC AACGCTTTTC AGATTAGAT TTTAAATTTT ATAAAAAAGC   | 3420 |
|    | ATCAGAAATA CCTGAGAGCG ACTTGGCTGA TTAGATATATA TTAGTTGGTT ATGATGCGGG | 3480 |
|    | TATCAATGAG GCATTTTTC GACGTTGCC GAATTTAAAA TGGATTGCAT GGTTTGCAAC    | 3540 |
| 45 | GGGTGTAAT ACATTGCCGT TAGATTATAT TGCAGATCAC GGCATACCTT TAACTAATGG   | 3600 |
|    | AAAAGGTGTT CAAGCTAAAC AATTATCTGA ATACATTTTA GCTTTCATTT TAGATGATTA  | 3660 |
| 50 | TAAAAAGATG AAACATCAT ATGATAACCA ACGACAACAT ATATATGATT CGAAAAAATC   | 3720 |
|    | TGGTAAACGC CTATCAGGAC AAACAGTTTT ATTTTTAGGT ACAGGTGCAA TTGCTACTAG  | 3780 |
|    | AACTCGAAG TTAGCAAAGG CTTTAAATAT GAATTTAATT GGTCTGAGCA AGTCAGTCA    | 3840 |
| 55 |  |      |

TGCTGACATT ATTATAAATG CTTTACCAGA AACGCAAGAA ACGATTCaTT TaCTAAAGAA 3960  
 AAAACATTTT GAATTAATGA AAGATGAAGC ACTTTTATA AATATAGGAC GAGGTAGCAT 4020  
 5 AGTTAAAGAA GCGCTCTTAA TAGAAGTATT AAAAAGTAAa GTTATTGAC ATGCATATTT 4080  
 AGATGTGTIT GAAATGAAC CTTTGAAACC TAATCATGAA TTATATGAAT TGGATAATGT 4140  
 10 AACTATAACA GGCATATAA CTGGTAATGA TTATGAAGCA AAGTATGACT TATTAGATAT 4200  
 TTTTAAAAAC AATCTAGTTA ATTTTCTCAA TAAGAATGGT CTAATTGAGA ATGAAGTTGa 4260  
 TGCTAAAAAA GGCATTATAA TGATATCATC ATGTAAATAT TGACACGCGC GCAATACTAC 4320  
 15 AGTTATATTT aTAGTAAgTl AATaATgATT ATATAAGaaA GATGGTGATA TAGATGAGTG 4380  
 TGTAAATAGA ATCAATTGAA CATGAAC TAGAATCAAT TGCATCATTG CGACAAGcAG 4440  
 GCGTAAGAAAT TACACCTCAA AGACAAGCAA TATTACGTTa TTTaATTTCt TCACATACTs 4500  
 20 ATCCAACaSc TGAATGaATT TATCAAGCAC TTTACCTGa TTTTCCAAAT ATAAGTGTTG 4560  
 CGACAATATA TAATAACTTA AGAGTGTTTA AAGATATTGG AATTGTAAAA GAATTACAT 4620  
 ATGGAGACTC ATCAAGTCGA TTCGACTTTa ATACACATAA TCATTATCAT ATTATATGTG 4680  
 25 AACAAATGTTG TAAGATTGTT GATTTTCAAT ATCCACAGTt AAATGAAATt GAAAGATTAG 4740  
 CTCAGCATAT GACTGACTTT GACGTAAAC ATCATCGAAT GGAAATTTAT GGAGTTTGTA 4800  
 30 AAGAAATGCCA AGATAAATAA TTTAACTTTG GTAGTATGAC AAATTAAAAA AGCGTTACTw 4860  
 ACTTCATATA AGTAAGCGTA ATATTTAAGA nGTTAAACGA CATGaAAGTt GTTTAACTTT 4920  
 TTT 4923

35 (2) INFORMATION FOR SEQ ID NO: 552:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 917 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 40 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 552:

45 TCCGCGCTTAA AAAACTTTTC CCAATTCAG CTGGGGCCTT TGGCATTAAT ATTAACCTCC 60  
 TGGTTCGGAT TAATTGGGAA CCTAACChTT TAGGCAATAA TTGGTTTAGG CAACTTCCAA 120  
 50 ATTGGTGGTT CAACCAACGT CTTTGGATAC CyTgcTCATT TAATTCTAAA ATGgTyrGAA 180  
 CGCATTTTGG TACCCAAAAA GgTGACGTTC GTTTCACGG TCTAATAAAT TGTCTAAGTT 240  
 GTGATTGGT TTCATTAATT GaaCACCATT TTGCTCTTCA GCAAGACCTT TGTCTACTTT 300

55

GCGATCCATT GTTGTGTCGA AAAGAATAGC ACCCAAAATT TTATCTGGTG AGAATGAAG 420  
 TGAAGTTACC ACACGTGTAC GCATATCGTG AACAAGTTGG AACATTTGCT CTTATTGCT 480  
 5 ATATTGATCT TCGTTTACAC CATATTCCTT AAGTGCTTTT GGTGTACTAC CACCACCTTG 540  
 GTCTAATGCG GCAATAAAGC CTTTCCATT TTTCAATTTT TCTAATTGCT CTTATTTCAT 600  
 10 ACTTCCACT CCTTAACCTT TCAATACACC TCCAGTATGA TAAAAATGAG AACATTTCTC 660  
 AAGTCATAAA CCTTGAAAGG TGTATAAAAT GTGAAAAATA ATTGTCAGTT TAATTAATAA 720  
 TATTATTTTA TTCTAGGTAT GACTAACGCC ATTAATGACA TAAAGAAAAT ATGTGTAATA 780  
 15 ATCCAAACGA TTAATTCGT CACACTAAAT TGAAAAATTG GACGTTGCGC AATAAATACT 840  
 AAAAAGGGAT ACAATGCTAT AAATAAGAAA AATAAGGGA TATAACATAG ATAGTAAAGC 900  
 CTTTGAAG TATGAAA 917

20 (2) INFORMATION FOR SEQ ID NO: 553:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 432 base pairs  
 (B) TYPE: nucleic acid  
 25 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 553:

ACTCCCATAT CCACGGCAA GATTTATAAT TTGTAATTCT AAAGTATAAT AGCCTATATT 60  
 TTTCAAATC TCTTTTTTA TGATTACAT CCTCTAAGT GCACCTAAAG GTTTTATGT 120  
 35 AATGTATTAA ATATTCCAT TTATACATAT ACATCTTGTA TAAAAAGAAA GAACCCATA 180  
 TATACTCAA AGGTATAGT AAGCTCTATA TATTATATAA AATTATTATA CTATTTGAA 240  
 AAACATAAGT AAACACTAA ACATGATTC TTTCACTTTT AGAAAACCTT TAATACTATA 300  
 40 AAAGACCCA CTCAGTCACT AGTTGGGCA GTTATTGTAT GCCTATTGAA CTCATGCGT 360  
 ATATTACAAT ACCTTTTCG CATATTCATA TAAGACTTG CATCTTAAAG CTTAATTGCT 420  
 ATCTCTTCT CT 432

45 (2) INFORMATION FOR SEQ ID NO: 554:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1374 base pairs  
 50 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

55

(2) INFORMATION FOR SEO ID NO: 555:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1472 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 555:



## EP 0 786 519 A2

TTACTCTAG GCATGTCCTT TTCACGGTIT GCTTATGATA ACGTTATCGA CATGTTAAAA 120  
 TTGTCCTITT TGAAAAATA ATTCTGGTGA GATAGACGCT ACTTGGAATT CaTcAGTTTG 180  
 5 TAATACGCA GtATAAtTAC CaTTACTAAA tTGaGTTAAT CGTTCaTATA AAGTACTAAT 240  
 AGGATAATAA ATGTTATCCG TTAAGCGCGC CGTATAGTTC ACTTGATACG TTTCGCCTTC 300  
 10 AACAAATGCT TGTTGGACAC GTTTAATATT AGTCATCATA ACCTCAGAAG ATTCAACAAA 360  
 TGAAAAATGA TACTTTGATA CATAAGAAGT TTGATGTCA TATGTTGAAT TTATGCTTTC 420  
 CGCTTTTCCA AAACATAAG CTGCTGCATA AATATCATCT TTAGCTAATG AATGTGTACA 480  
 15 CATAGCATGA TTAAAAACT TTGCCGCTTC GTAACCTAAA TATAACGAGA CATATCTACC 540  
 TTGTCGtKt GCGCTTGtGC AAAGTGTATC ACTTCTCCCA CATCAGCCAA CTTAGTAGCA 600  
 ACATACTTCT TTATAAATCC CTTTAATTGA ATATGGTATT GCTTATATTC ATTTTCAGTT 660  
 20 AAATAGTAGC GATAATTATA TTCTATTCTC ACAGTAATCA CCTACCTTCG ATAAAAATAA 720  
 TTCAACTTGT CGATAACCGT ATTCACCTCA AATAGATTCA GGATGATATT GCACACCAAA 780  
 AACCGGAAAT CTAATATGCT CAAATGCCAT AATAATCGCT TCATCGTTTT TTGCTGTAAT 840  
 25 CTTTAAGCAA TTGGAAAAG TCGCTCGTTC AGCAATTAAT GAATGATAAC GCATTACATT 900  
 GAAATTTTGA GGCAGTCCTT GAAAAATACC TTCATTGGTA TGGCGTAACT GTGTAGTATG 960  
 TCCGTGTACA GGATGATAGC CGTGAATGAT ATTTCCACCA AAATAAGACA CGATACATTG 1020  
 30 AAATCCTAAA CATACACCTA GTATAGGTAC ACGCTGATAA AATTGTTCTA ACACCTTCATT 1080  
 CAAGATAGGA TAATCATCCG GATTACCCGG CCCAGGCGAA ATAACAATTG CTTTGGCTT 1140  
 35 CATATTAATG ACGCTTCTA TCAGCAGATT ATCAATACCA ACAACTTGAA CTGTAGTTT 1200  
 CGTTTGAGTC TTAATATAGT CTATTAAATT ATATGTAAT GaATCATTAT TATCTATGAC 1260  
 TAGAATCAIT GTATACTCCG TTCTAAATGT GTTTATTTT TATAATATGT ATTGGATGTA 1320  
 40 GCTAAACTT TAAAGCATT GTCAATATCC TGACCTTGAT TTTAACTAAT ATATGGTATA 1380  
 TTCTATTCAT CGTACATAAA TGAATATCAG AGGTCTCTAG CTGAAACCCCT CTATAAAAAA 1440  
 CTAGGCCAIT GAAATTTCAA ACATTGTTG GG 1472

(2) INFORMATION FOR SEQ ID NO: 556:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1054 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

EP 0 786 519 A2

AGAACAGCAA GGATTACTTCTA CTGAGGAATT AAAGAAGGAT ATTTTAAAC AGAACAAATT 60  
 ACAAAGTGTG GAAGACCTAT ATAGGCCTTT TAAACAAAG AAAAAGACAA GGGCAACTGA 120  
 5 GCGAAGACGT AAAGGGTTAG AGCCATTAGC GATATGGATG AAGCCACGTA AACATGAAGT 180  
 CTCAATTGAA GAAAAGCAC AACAAATTAT AAATGAAGAA GTGCAATCGG TTGAAGATGC 240  
 10 TATCAAAGGT GCACAAGATA TTATTGCGGA ACAAATTCTA GATAATCCTA AATATAGAAC 300  
 AAAAATTTTA AAAGATATGT ATCATCAAGG TGTGTAACT ACATCTAAAA AGAAAAATGC 360  
 TGAAGATGAA AAAGGTATTT TTGAAATGTA CTATGCATAT AGTGAGCCAA TTAACGCAT 420  
 15 TGCTAATCAT AGAGTTTATG CTGTTAATCG TGGTGAAAA GAGAAAGTAT TATCTGTAAA 480  
 GTTGAATTC GATACGACAT CAGTAGAGGA TTTCTATTGCA CGTCAAGAAA TCAATCATAA 540  
 TAATGTAAAT CGCAGTTATA TTTTAGAGGC GATTAAAGAT AGCTTGAAAC GCTTAATTGT 600  
 20 CCCTTCGATA GAGCGTGAAA TCCATGCTGA TTTAACGTAA AAAGCTGAAA ATCATGCAAT 660  
 AGATGTTTTT AGTGAAACT TAAGAAATCT ATTTACTGCA CCTCCAATGA AAGGTAAACA 720  
 AATATTAGC GTAGATCCAG CATTTAGAAC AGTTTGTA TTAGCAGTCA TTAACCCATT 780  
 25 CGGTACTTTT ATAGCAAAAG GTGTGATTTA TCCGCATCCA CCAGTTTCTA AAAAAGAGGC 840  
 AGCAGAGAAG GATTTGTAC AAATGGTTAA AGCGTATGAT GTGCAATTAA TTGCAATTGG 900  
 30 CAATGGTACT GCAAGTCGTG AAACAGAAC AATTGTTGCA GATTTAATTA AAAAGCATCA 960  
 GTTGCCAGTA CAATTATCA TTGTCAATGA AGCGGGCGCT TCAGTATACT CAGCATCAGA 1020  
 AATTGCTAGA GATGAATTC CTGATTTTCA AGTG 1054  
 35 (2) INFORMATION FOR SEQ ID NO: 557:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1057 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 40 (D) TOPOLOGY: linear  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 557:  
 45 AAATGTCAGA ATACAAGAAA AAAATAATTG AATTAAATGA AAGTAATTTA ACAGGATATG 60  
 AAATTTCTAA AAAAAGTGA GTTCTCAAT ACGTACTTTC ACAATTAAGA CAGGGCAAC 120  
 50 GCGAAGTAGA TAATCTAACC CTGAATACAA CAGAAAAATT ATATGAATAT GCCAATAAAG 180  
 TTTTGTAATT TAACTAATGT ATAAATTAAT CAAGCTATGT TTATTTGATT TAACTATTAA 240  
 TAAAAATCAT ATGGTGAATG GATATTATAA TAATTTAAAT ACAAAAATAG TAGATTCCAA 300  
 55

EP 0 786 519 A2

AAGGGAAAAA AAGTGTTAAG TTTTAAATGA TAAAAAGAT TGGAATGGAT CGTCTTGAAA 420  
 TGCCTCCCTC AAAGTTTTC A TTTTTCAT GTCGACTTCG AAGGGGGCAT TTTTATTAAA 480  
 5 TTGTTATAGC TTTTATATT TGTATAATGA ACATATAAGT TTAAGAAGGT GCGAGTGAAG 540  
 GAAATAAAAA AGCTCAAATG TACCAAAATG TTAATCTTAA TAAATCTCTA CTTTATAAAG 600  
 10 ATTGAATGGA CATTGAGCG TTAATCAGTC AGGAGGGACT TTCCCTCCTA CAATTTAATA 660  
 ATAATACTTG CTTCACCACT ATACAAGGAG TGAGTTGTTA TGTTCAAAGT GAATTATTTCG 720  
 ATTTTAAGTT ATTATCCAGA ATATAATATC GCAGTAAGTT GGCAACGTTT AAGAGAAGGA 780  
 15 AAAACAATAA AAAACAAGAT TTAATACTGC TGCCTCATGA GCGCTTGAA CATTATTGTA 840  
 TGAATAAGTA TAATTTCAAC TATGATTATG CACATAAAAT TGTATCAAAA AAATACGATT 900  
 ATTCAATTTT TATAAAAAAG AAGGTGGATT AAATGCTTAC ATTAATAAAA TTGGAAGAA 960  
 20 GATGaACAGG tTATAATATA TGrATATATA CCTGaAGATG aTATAAGTAC CGGGTAAAGG 1020  
 GTCCCGTACC TTTTAAATTA AAAAAGTTCC AGGGGGT 1057

(2) INFORMATION FOR SEQ ID NO: 558:

25 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 3754 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 30 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 558:

35 CTGTGCTGTA TTACtTTAA TTGACGACA TTGAGAAGT ATTATGATGG rTGtAGTTAT 60  
 TTTATGAAAG TAATGTATTA ACAATCGAAT TACTAAGTCT ATATTGGTAG GAATATCGAT 120  
 GTTTAGTTTA AATGGAATGC ACTATTTAAG TTTTAAATAT GGAGATGTTT GTGACTTTTC 180  
 40 GATGATTAAg ATTTTtATAG GTGTGCATCA TTTCAATAAA ACTTTGTGTT TAAAATTGAG 240  
 CTTAGGAAAT CGATAGGTTT AGATGAGGAT ATTGTtGAAG TTATGTGTCT TGTATCCTTA 300  
 GTTGTtATAA AAGCGCAAAA AATAGCACCG CTTTCTCTTT ATCTGTGTAG AAAGGATGCT 360  
 45 ATTATTGTAA AACAATAGTT TTAATTTtAT TTTCTGAtAT ATCATATGTr ATTtACCTG 420  
 TATCAATTTT TATCGAATTA TAACCATCAA AATTATCAAC TTTATCATCA AAGTCTATCA 480  
 CTTTCCCAAT TAATATTTTA TTATTAGTAA GCGTTAATTT GACTAATTTG CCTATGTATG 540  
 50 ACTGTAAAT CATATTTAAT CACTCCTTTT TAATATACGG AACTACATGA AAACGAGTTT 600  
 TAGAATAATG AACCTTACCT AATTTCGTTT CAATATATTT ACCATTGACA TAAGATTTAC 660

55

EP 0 786 519 A2

|    |  |      |
|----|--|------|
|    | AAACTGGAAT ATCAATAAAT TCTTTTTIAG TCATCTTTTC ACTTTCATTA GAATCTATTA  | 780  |
|    | TAGTATAACT TGGTAATAAA GATGAATTAT TTTTCTTATT ATTATATTCA ITTGTGCATA  | 840  |
| 5  | AAATATGACG CTITTTGCTTA ACATTATTCA ATTTCATCGT CATTTTACCA TTACTTATCA | 900  |
|    | TTTCAAGCAT TTCTTTTTTA GCTTTTTGTA GAAAGGATGC TATTTTGTA TTTTTTAAT    | 960  |
| 10 | AATCTAATAC TTTTATCTC AGTTTCATCA AACGAATAAA CAGCAAAATC AGTCTGTATA   | 1020 |
|    | TCTACAACAA GATTTCCTGT TTCGCTCTCG AATTCAATTT CATAGTCAGT TACAACCCCT  | 1080 |
|    | TCGTATTCTT CATTATTCAA AAGCGTGATA ATAACATCTT TACGATATGC ATCTmCAATT  | 1140 |
| 15 | CTCAATACTT TTCACCTACT TATCAATATA AGGTACTATA TGAGCACCTG GTCTTCGAAT  | 1200 |
|    | AATGGCACIT CCCTTTCTGG TTCAATATA CACATTTTCG ATATGTATTT TTCCAATAT    | 1260 |
|    | TTGATTAAAA TTAATAATCT CTTTCAAATC AAATCGCTCA TCACCTTAATA TTAGATTGCC | 1320 |
| 20 | TGTTGACATT TTTTCTCTTA ACAATTCATT CAATAAATCT ATAGAAAGTA TTGTATAGCT  | 1380 |
|    | AGGCAATTTT TTATTATTTA AAATGGCTCT TTTTTTATTT TCATTATATA GGTGATGACC  | 1440 |
|    | TAACATATAT CTATTTTATT TTCCGCAATT TATTTCTATT TTTATTTTAC CATTTTAAAT  | 1500 |
| 25 | CATTTCTTTC ATCTGGTTTT TAGCTTTTTTC CTGtAATTAT GCTTCTTTTA CTCTACTTG  | 1560 |
|    | ATATTTACCT TCACGCTCTT TAAAGAACTT GTCCGCGCAA TTGCCAACAT GTGGCACTGT  | 1620 |
| 30 | GGTACTTCTA CACCAAGGAT GCATAGGTGG CGCATTCACT CCTGGTATCA TATCTTTAAC  | 1680 |
|    | TTTAAATATT TTTCGGTTAA GTGAATGACA TAAITTAGAT GTTTTACTAT CTATTTTGGC  | 1740 |
|    | AACATATTTA TATTCGCCAT CTTCACCAAG TTCTTTTAAA TATGTTAACT TTTTGCTTC   | 1800 |
| 35 | TGCATTTTCA GTAAATAGTT AAAAAGCGT ATAAAAATAG CACCACCTTC TCTTTAkCTG   | 1860 |
|    | TCTAAAAGG ATGCTATTTA TCTTTTGAAT TTGAATTCCT TTTGCTTTT TCTATACTTT    | 1920 |
|    | CAAATTCCTC AACTAATTC TTAAGAAT CACTCAATTC TTTTGCAATT ACATTTCCAT     | 1980 |
| 40 | CTAATTGTGA ATCTAACATA ATTAAAAATCA TCTCACTTTA TATTTAATCa TATTTTACT  | 2040 |
|    | ATAAGTTTT TTCAATAAAT TTCAATATG GCTATCATTa ATGATATTAA TATGCGTAAA    | 2100 |
| 45 | ATATTTAGCA CAAAATTAC TCACTATTTT ACCATGAAAC CTATTTGACT TGGTAATAAA   | 2160 |
|    | TTTTACTTGT CCCTTATTAG TAACGATTGT CATTGATTTT ATTGATGGAT GCTTAAAAAA  | 2220 |
|    | TGTAAATAAA TCATATCTG AAAATCCTGA CTGTCCAGGA TGGTTATGTA ACATAACAAT   | 2280 |
| 50 | TGAATTCGGT TTAGCTTTAA ATAATAATTC GGTTCCTTGT TACCCTGGCA CAAAAGATAC  | 2340 |
|    | ACTATCTTGA TTGACATATA CTTTGTGAAA TTTACCATCT TTTAACAAAT AAGCTACTTC  | 2400 |
| 55 | ATTGCTATCA TTGTTTCCA TGAGATATAC CTCCGTTTAT AGtCTGTCGA CTGATATTCC   | 2460 |

## EP 0 786 519 A2

|    |  |      |
|----|--|------|
|    | GAAGGTTAAA ATTTTAAAG TAAACTTTCT TTATCCCAT GCTACGAGTT CAGATTGAGG    | 2580 |
|    | AAATAGCCCT TTACTAGTAT TTATGTATAT TCTGTCTATG GCATGAATAA AATAATTATC  | 2640 |
| 5  | TCTTGATATT TTTTCTAAAC TAGATTTTTT AGCATTGATA ACTTCAAGAC TATCTATATC  | 2700 |
|    | CATTTGAATA ATACCAGGCT TAATATTTTC ATCATTATTA GGAAAAATTT TATATGTAAAC | 2760 |
|    | ACTTTTATCA TTAATTTCTT TTATTTTAA TATTAGCAAT CATTTCCACC TCTAATTAAT   | 2820 |
| 10 | TAAAACTA TAATTATAT TTATTTCTGT AAGTTTATGT GCCTCTATAT AGTGTAATTT     | 2880 |
|    | ATACTTATTC ATTAGATAGT GTTCAAGAGC TTCATGTTTC TACATTATTA TATCCATTTT  | 2940 |
| 15 | TTTAATATT TTCCCTTCTC TTAACGTTG CCAACTTTGA GCCATATAAA AGTCAGGATC    | 3000 |
|    | AAATGTTTA AATCCACTTT CTAATAAATA CTTATTTTCA AATATATGTT CATAACTCT    | 3060 |
|    | TTGAATTAAA TTTTATTTA TATTAGTATT TTAGCAATT TTAGAAATCT CTATCTGTTT    | 3120 |
| 20 | ATCTCGATTT CTAACGAAAT TATAATAAAT TTGAGCATGT CTGTTCTTTT TGATACCGTA  | 3180 |
|    | TTCATCACTT TTATTTTAA GTGCACCTGA TTCAATAAAA CAACCTTCTA CTTGATATTT   | 3240 |
|    | ACCTTCACGC TCTTTAAAGA ACTTGTCTCG CCAATTGCCG ACATGTGGCA CTGTGGTACT  | 3300 |
| 25 | TCTACACCAA GGATGCATAG GTGGCGCATT CACACCTGGT ATCATATCTT TAACTTTAAA  | 3360 |
|    | TATTTTCCG TTGAGTGAAT GACATAATTT AGATGTTTTA CTATCTATTT TGGCAACATA   | 3420 |
| 30 | TTTATATTCG CCATCTTCCA CCAAGTTCyT TTAATATATG TAACTTTTGT GACTCTCTyT  | 3480 |
|    | TTTCAACGAA TAATGAAAAA AGCATATAAA AATAGCATCG CTTTCTCTTT ATCTGTGTAG  | 3540 |
|    | AAAGGATGCT TTAATACCaT GCTATTTTAT AATTTTCGGG AAATTCCTGC TTCTCGATAA  | 3600 |
| 35 | AGTCTCTTAC TACAGAAAAA GACTTATTAC GATATAACAT AAAATATCTT TCATTTTCTA  | 3660 |
|    | TTTCAGATAA TGAGAAATCT ATAACCTCTG CATCTTTTTT ATTTAAAGTT ACTGAACCTT  | 3720 |
|    | TACCGTTACT TATATCATCT TCAGGTATAT ATTC                              | 3754 |
| 40 | (2) INFORMATION FOR SEQ ID NO: 559:                                |      |
|    | (i) SEQUENCE CHARACTERISTICS:                                      |      |
|    | (A) LENGTH: 815 base pairs   |      |
|    | (B) TYPE: nucleic acid   |      |
| 45 | (C) STRANDEDNESS: double   |      |
|    | (D) TOPOLOGY: linear   |      |
|    | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 559:                         |      |
| 50 | ATTTAACTAA AACTATAaAT AATCAAATGA TATTGGAAGA TATTAGCATA GATATCGAAA  | 60   |
|    | AAGGTAAATT GACTTCTTTA ATTGGACCTA ATGGTGCGGG TAAGAGTACT TTACTTTTCAG | 120  |
| 55 |  |      |

EP 0 786 519 A2

|    |            |             |             |            |            |            |     |
|----|------------|-------------|-------------|------------|------------|------------|-----|
| 5  | CTGATTATTA | AAAAAATGAC  | TTGTGCAAAA  | AAATATCTAT | ATTAAAACAA | ACAAACCATA | 240 |
|    | CTGAAATGAA | TATTACGGTA  | GAGCAGTTGG  | TAAACTTTGG | ACGATTCCCT | TATTCTAAAG | 300 |
|    | GTCGTTTGAC | GAAAGAGGAT  | CATGATATTG  | TCAAATGATG | GCTAGATTTG | TTGCAACTAC | 360 |
| 10 | AAGATATCAG | AAATCTGTAAT | ATTAAGTCAT  | TATCTGGTGG | ACAAAGTCAG | CGTGCATACA | 420 |
|    | TTGCAATGAC | AATAGCACAA  | GATACTGAAT  | ATATTTTGCT | AGATGAACCA | TTAAATAAAT | 480 |
|    | TAGATATGAA | GCATGCTGTT  | CAAAATTATGC | AAACGTTAAA | AATGTTAGCG | CATAAAATGA | 540 |
| 15 | ATAAAGCGAT | TGTCATTGTG  | TTACATGATA  | TTAACTTTGC | GCTCCTGTAT | TCAGATCAGA | 600 |
|    | TTGTAGCAAT | GAAAAACGGA  | CAACTAGTTA  | AGTCAGATTT | GAAAGATAAT | GTCAATCAAA | 660 |
|    | GTAGTGTGTT | AAGTGATTTA  | TATGACATGA  | ATATTCAAAT | TGAACATATA | AGAAATCAAA | 720 |
| 20 | GGATTGTGTT | ATATTTTAAG  | GATGTGATAA  | TTGGAGACAC | TTTAAAGGGG | TGATGCGCCA | 780 |
|    | ATTAAAGAAG | GGTAAACGAT  | AAAGCATTTTA | TTTAT      |            |            | 815 |

(2) INFORMATION FOR SEQ ID NO: 560:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 919 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 560:

|    |             |            |             |             |             |             |     |
|----|-------------|------------|-------------|-------------|-------------|-------------|-----|
|    | GAACGAAATA  | ATAAAATTAC | TGAGTTATTA  | GTGAAAAAG   | CTAATAAAC   | TGATGATGT   | 60  |
| 38 | CTCGATAAGT  | TGATTAAAT  | TTTAAATATA  | CGCATACACA  | CTACTAAAA   | TGATTATTATA | 120 |
|    | TTATGTTTAA  | TGAGCCTC   | CTAAATTTGC  | AAGATAGACA  | ATTTAGGAGG  | CGTGTTTATT  | 180 |
|    | TTTATTGAGC  | TCTAACTATA | AAGATATATA  | ATTGACATT   | TACAAATGAT  | GTAATAACG   | 240 |
| 40 | CAATTCTCAT  | CATCGCTGTA | ACAATTCAT   | GGTTTAATAT  | GCAATGAGCA  | TATACTTTTT  | 300 |
|    | AAATAGTATT  | ATTCACTAGT | TTTAACAATC  | AATTAAATGG  | TATATGATAC  | TTTTATTGGT  | 360 |
|    | TATTTTTIATC | CCATAGTGTG | ATAATTTACTA | TTTTTTCATTC | ATAATAAAGG  | TTTAAAGCAT  | 420 |
| 45 | GTTAATAGTG  | TGTAAGATTA | ACATGTACTG  | AAAAACATGT  | TTAACAATAT  | GATATAAGGA  | 480 |
|    | GTGACGTACA  | TGATCCGCTT | AGGTAATAATG | TCAGATTTAG  | ATCAAACTTT  | AAATCTAGTA  | 540 |
|    | GAAGAAGCAA  | AAGAATTAAT | GAAAGAACAAC | GACAACGAGC  | AATGTGGACGA | TCAGACCCCA  | 600 |
|    | CTTTTAGAAC  | ATTTTGAAGA | AGATATTGCT  | AAAGATTATT  | TGTACGTTAT  | AGAGGAAAAAT | 660 |
| 50 | GACAAAATTT  | ATGGCTTTAT | TGTTGTGCAC  | GACACCAAG   | CAGAATGGTA  | YGHYGACATT  | 720 |

## EP 0 786 519 A2

TATAAAGGAG CTGCTACAGA ATTATTCAAT TATGTTATTG ATGTAGITAA AGCACGTGGT 840  
 GCAGAAGTTA TTTTAAACGA CACCTTTGCG TTAACAACAC CTGCACAAGG TTTATTTGCC 900  
 5 AAATTTGGAT TTCATAAGG 919

(2) INFORMATION FOR SEQ ID NO: 561:

(i) SEQUENCE CHARACTERISTICS:  
 10 (A) LENGTH: 518 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 561:

ATCATATAAA CCGGCTGCTT CTAACAAACGA AAATACTGTT ACTGGACCTA AAAATTTTAA 60  
 20 CCCGTTATTG TTTAAATCTT TAGATAGTTG TGTTCAGTA TCATCAACTG TGATACGATC 120  
 AGAAGCATGT TCATACTGCA AATCTTTAGG CTTACCATTT ACATATGACC ATAAAAATTT 180  
 ACTAAACTA CCATATGCTT GTTCAATTTT TAAATACCCT TGAGCTTGAT TAACAATTGC 240  
 25 TTCTAATTTT TTACGATGAT GAACGATATT TGGAAAAAGTC ATTAAAGCGT CGATATCTTG 300  
 AGCGGTCATT TGTGCTACCT TTTCTGGTTC GAAATCATAA ATGCTTCTTT CATAGGCTTC 360  
 TTTCTTTTIT AAAATAGTTA ACCAAGATAG CCCAGCATGT TGTGATTCTA ATGCTAAAAA 420  
 30 TTTAAACAAT GCCTTGCAAT CATAGAGCGG TTGTCCCAT ACATGGATCm TGATAGTCTA 480  
 AGTAGACTGG GATCTTTAGT ACCAAATGCG CATTCATT 518

(2) INFORMATION FOR SEQ ID NO: 562:

(i) SEQUENCE CHARACTERISTICS:  
 35 (A) LENGTH: 1539 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 40 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 562:

45 CTTATTTTAA AGTAATCAA TCAATGTATT ATAATCCGAA TAGTCCGCAT AAAGCTGGTT 60  
 TGCAGGCAAA tCAATTACTA CAACAAGCAA AAACCCAAAT TAATGCAATG rTTaATTCAA 120  
 50 AAACAAATTA TGATGTTGTA TTCACTAGTG GTGCAcTGAA TCCAATAATC TTGCTTTTAA 180  
 AGGTATTGCC TATCGTAAAT TTGATACAGC GAAGGAAATA ATTACATCCG TGTTAGAGCA 240  
 TCCGTCCGTA TTAGAGGTTG TAAGATATTT GGAAGCACAC GAAGGATTTA AAGTTAAATA 300

55

|   |            |             |            |      |
|---|------------|-------------|------------|------|
| A | AAATAAGTGA | ACTGGACAAA  | TACAGCCTAT | 420  |
| A | TCCTAAGGCA | CATTTTTCATG | TAGATGCGGT | 480  |
| T | CAATAACATA | GATAGTATTA  | GTTTAAAGTG | 540  |
| G | CGCTTCTACT | GTAATACACA  | TTCAAAATGT | 600  |
| A | ATATGGTGTT | AGAAGTGGAA  | CAGTTAAATT | 660  |
| T | GAGTAGTAGT | AATGAAAAC   | TTGAAGCATT | 720  |
| A | CGTCGCTGCA | TTTTTAAATA  | AATATCATGG | 780  |
| C | ACCAATCGTT | TTAAATAATTA | TTTCTCTGG  | 840  |
| T | TTCAAAATAT | GACATTATGA  | TATCTACGAC | 900  |
| T | AAATGAAGTA | TTGCGTCGAA  | TGGGATTATC | 960  |
| T | ATCATTTGGG | GCTACTACAA  | CTAAAGAAGA | 1020 |
| T | CATTTATGAG | GAAATTAAAG  | AGTTGCTAAA | 1080 |
| A | TACGGGGAGT | TAACATTAAA  | GGGTTCAAAT | 1140 |
| T | AATGTAAATA | AGTCaTTAAA  | AGGACTTGAT | 1200 |
| T | ATGTATATTG | AACTTGAAGA  | CCATGCaGAT | 1260 |
| A | ATTTTCGGTA | TTAAATCTAT  | TAGTCCAGTA | 1320 |
| A | AGTGcAGCGG | CAATTAAATT  | gCGCAGaATT | 1380 |
| T | GAGCGGTGCC | CGaTTAAAA   | TTCCCCAATT | 1440 |
| G | GGGTGGTGCC | AGTATTGGAG  | CACCTGGCCA | 1500 |
| C | CCAGGAATT  |             |            | 1539 |

(2) INFORMATION FOR SEQ ID NO: 563:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 968 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 563:

|             |            |            |            |            |            |     |
|-------------|------------|------------|------------|------------|------------|-----|
| ATACGCTATA  | CATGTGTTCT | TTAAAAATGT | GATAAGGAGT | TTAGGATGGT | TTATTTAAAA | 60  |
| TCAATAGATG  | CCATTGGATT | TAAGTCTTTT | GCAGATCAAA | CCAATGTTCA | ATTCGATAAA | 120 |
| GGTGTAAC TG | CAATTGTTGG | TCCAAATGGA | AGCGGTAAAA | GTAATATTAC | AGATGCTATT | 180 |





## EP 0 786 519 A2

(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 565:

|  |      |
|--|------|
| ATGTTTTGTC AATATCAATT GTTGGAGATA AATCCGCTTG TATAAACGA ACTTTATCAT   | 60   |
| CATTAAATTT GCTAGTTAAT TCATTTATAT CAGTACGATA ATATTGCACA TAAACTTCAA  | 120  |
| ATCCATCAGT TAATAATTGT TTGACTATCT CAGAACCAAT TGAACCAGAA CCACCTAATA  | 180  |
| CTAATGCTTT CATTACTTTT TAATCTCCAA ACGACTATCG ACTTGCTGAT CTAATTTTAA  | 240  |
| ATATAACGAT GACGTTTCGT TAATACTATC TAACTGATA TTTTCAACAA TGCTTAACAT   | 300  |
| GTCAACACAG CTAACACCTT CAAAATACAA TTTAGTATAT TGATTAGCAA TATATTTCAGG | 360  |
| TGAGTTTAAA CTTGATATGA ATTCACCTAT AAATGCTTTT TTCAAAAGTT CAAATGCTTC  | 420  |
| TGCATCTGG AAAATGCTTT TTTTATCACG CAACTCATCT AATAATAATT TTTTAAATTT   | 480  |
| ATCTGTTTCT TCAGTAGCAC TTGTACGAT TGA AAAACTA TACGTCGGCT CTAGTACAAA  | 540  |
| TTGATAACCA AATGTATCAT CGATAAGTCC TTCGTTTAAT AAATTCGTAT AAAAATCTGT  | 600  |
| TTCTTCCCCA AAAATTAAC TAAAGAATAA TGACATTCTT AAATCAGTT GTACATATTT    | 660  |
| TTGAGGCGCT TCTTGTAAAT GTTTATTTTT AAAACCAAGC ATTAGTCTTG GTGATTGAAT  | 720  |
| TTTCATAGAT TCAGTAACAA ATGCTTCTTT AACATCCTCC GGTTCATCAA CAAGTCCTCG  | 780  |
| TTGATTTTGG GGTGGTTAA CTTTATTACG AGCATCCTCG TGTGTTTTTA CTATTCgACA   | 840  |
| TATTGCTTCA GGATCCACAT CGCCAACAAC AAATAAAACC ATATTGTATG GATGATAAAA  | 900  |
| CGTTTCATAA CATAGATACA AATCATCTTT TGTAATATCG TATATACCTT CTACACTACC  | 960  |
| GGCAATATCA ACACGTATTG GATGTTGTTG ATACATTGCA CGCAALGTAT TaaACATTAA  | 1020 |
| TTTATATCCA GGTGTCTCTT GaTACATTTT TATTTCTTCT GCAATAATAC CTTTTTCTTT  | 1080 |
| ATCAACAGTT TCTTTTGTA AATAAGGCGT TTCaACCATT GTAAGTAAAC GTTTAATGTT   | 1140 |
| GTTTTCAATA TtATCAGTTG CACTGAACAA GTAGCTTGTA CGATCAAAGC TTGtAAACGC  | 1200 |
| ATTTGCTTGT GCGTTATCTT CAGCAAACGC AGTAAATAAG cTCTTCTTTC TTTTTCAAAT  | 1260 |
| AATTTATGTT CcTAAAAAGT GAGCAACTCC ATCAGGTACA GTACAAATG GGTCTTGTCC   | 1320 |
| AAGGGGTTTG AATTGATTAT CTAATGAACC AAATTGTGTA GTGTAAGTGA CAAATGCTCT  | 1380 |
| TTGAAAAACCT GGTtKGGGGA TAATAAATAA TCGTAAACCA TTTTCTAATC CTTGTTCGAA | 1440 |
| tACTCTTTG TCTATTAATT CATAATAACG CTCTTCATT ATTTATCCCC TCCTTTTGTC    | 1500 |
| AACACATAGA TtGTATCTAA AAATGCTTTT TCAGCAACAG AAACAATATC TTCGCACTT   | 1560 |

55

EP 0 786 519 A2

TTATGCATAA TCTCTATAAT ACTTTTCGGA CGATCTTCAG ATTCATATCG ATGAGAAATG 1680  
 ATTACTTTTT TAGCTAACTC TAATTTTTCT TCAGTGAAAT CTCCTGCTTT TATTTTTTCA 1740  
 5 AATTCACCTA TAATAGTGTC TTTTGCAGTT TCGTACTTAT CACTTGAAC CCCTACTCAA 1800  
 ACAATAAAT AGCCATTTTT GCCATCAATT TGTGAATGTA TAGAGTACGC TAAACTTTGC 1860  
 10 TTTTCTCGCA CTTCATTAAA TAAACAGAT GAAGGATCTC CTCCAAACAT CATGTTAAAT 1920  
 ACAACAAAGG CAGCATATCC ACTTTGTCCA TATTGTGTTG GAAAACGGTA TCCCATATTT 1980  
 AATTTAGCTT GATCCAGTC ATCATATTCA ACAATAAAT CACTTCTTC ATCGTGTAAT 2040  
 15 TGATGAGTAG AATGTTGGAA TTGATGTTTA TCGAATGGTT TAAGTGCAAA TTTTTCACGT 2100  
 ATTTGTTTCT CAACACTTTC AGGTTCTACA TTGCCGACAA CATAAACAGA ACATTGATCA 2160  
 TTATTAATCA TTGATTGATA TGTATGATAT AGTGTTTCAG CAGTAATATG TGGGATTGT 2220  
 20 TCTAGTTGTC CTGTAGATAA GTATTTATAT GCTTCAATTT CAAACATATG GTCGAGTAAT 2280  
 TTTAAAAACG AATATTGTGC TTTATTATCT ACCATTGCTT CTATTTTTTT GGCTAATAAT 2340  
 GTTTTCTCTT GGTTAACAAA ATTATCAITG AATGCTTTAT TTTCAATTAA TGGATTCCAA 2400  
 25 ATGATTTCTT GtAATAAATC TAATCCTTGA TTAATAAATG AwTACCCKT TCyTAAATAA 2460  
 CGKkCaTTAA caATTyCTAA tGaaaATGtA ATgAcATGCT GaTCTTTGAA TTTTGAAAT 2520  
 30 GTACTATTCA CATACGCACC ATATAAATCG GCTA 2554

(2) INFORMATION FOR SEQ ID NO: 566:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1424 base pairs  
 35 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 566:

TnTCGTATT TThAATTGGT TTTGTATGGT TTAAGTTATA TCAATATACA ACAnACCCTA 60  
 45 AAGCTGATAT CCGAGTATC AITTTTASTA CGATTGGTTT TGGTCTTTG TTATATGGTT 120  
 TCTCAGAAGC TGGCAACAAA GGTGCGGTT CAGTAGAGAT AGAAACAATG TTTGCGATTG 180  
 GTATTATCTT TATTATTCTA TTCTTATTA GAGAATTAAG AATGAAATCA CCAATGTTGA 240  
 50 AITTTAGAAGT ATTGAAATTC CcAACATTTA CATTAAACAA AATTATTAAT ATGGTTGTAA 300  
 TGTTAAGTTT ATATGGTGGT ATGATTTTAT TACCGATTTA TTACAAAAAT TTACGCGGAT 360  
 55 TcTCAGCATT AGATTCCGGA TTGTTATTAT TACCTGGTTC TCTAATTATG GGTCTACTAG 420

## EP 0 786 519 A2

TTGCTGTAAT GACTTATGCA ACATGGGAAT TAACTAAATT AAATATGGAT ACACCATATA 540  
 TGACAATCAT GGGTATCTAT GTACTTCGTT CATTTGGTAT GGCATTTATA ATGATGCCAA 600  
 5 TGGTAACTGC AGCTATTAAT GCGTTACCGG GACGACTTGC CTCTCATGGT AATGCTTTCT 660  
 TAAATACGAT GCGTCAATTA GCAGGCTCTA TAGGTACAGC AATCTTAGTT ACTGTAATGA 720  
 10 CAACACAAC TACACAACAC TTATCAGCTT TTGGGGAAGA GTTAGATAAA ACGAATCCTG 780  
 TTGTACAAGA TCATATGCGT GAATTAGCAT CACAATATGG CGGACAAGAA GGCACAATGA 840  
 AAGTGTACT ACAATTGTGA AATAAAGTAG CAACGGTTGA AGGTATTAAT GATGCATTTA 900  
 15 TAGTTGCAAC GATATTTAGC ATCATGCGCT TAATTTTATG TTTATTTTAA CAAAGTAATA 960  
 AAAAAGCAAA AGCTACAGCT CAAAAGTTAG ATGCAGATAA TAGTATCAAT CATGAATAAA 1020  
 TAAAAATAAT TAATTGAAGT GTGACTAATC AAAAATTATG TTGTGGGAC ATGATTTTAA 1080  
 20 AAGTATCGGT GCCAAATATG GTTATCGATA CTTTTTTTAT TTGTTGATTT ATAGAATGTT 1140  
 AGAGGAATTA TATTAATAAT TGGCATTGAC GTAGTAGGTC ATTAATAAAG AAAAAGCAGG 1200  
 AAGTGGGTCA ACGAAATGAA TTTTGTGAAA ATAACAATTG TGTCCEAATC CCTACTATAT 1260  
 25 AACATTATTT TAAACGAGGC ATGCGATTAC GGAAGAATAA GCTTATAACA AGTAAACCGA 1320  
 TGCTACAGCC AAGTAAAATG ATGCCGTTAT GAATAGCGTC ACTTgCTGTA ATCACTTGAT 1380  
 30 CTGTTGGTAC ATTTAAATAA TATTTTTTGA AAACATCTGC AATT 1424

(2) INFORMATION FOR SEQ ID NO: 567:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 676 base pairs  
 35 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 567:

TCTCCCATGT TCTGCTAAAT GACGCATCAC TTTTACTTCA TGAGGGCGTCA ATACACGTCC 60  
 45 TTCACCAGCA TTCAAACCGA CAACATTTAA AGGCCCATAT TCAATACGAG ACAGTTTCGT 120  
 CACTTGATGA CCAAAATGTT CGAACATTCT TCTGACTTGG CGATTACGAC CTTCGTGAAT 180  
 TGTAAATGCA ACCAATGTG TGTTTTATC TTTATCTTGT TTCTTAACTT TCACCTTCAGC 240  
 50 CGGTTGCGTC ATACCATCTT CTAATTCAAT ACCTTTTCT AGCGCTTCCA CTTCCTTCTCT 300  
 CATTAAATAA CCTTTTAATT TCGAACATA TTTTTCCTTA ATTIGATATC TTGGATGTGT 360  
 CATTAAATTA GTAAATTCAC CATCATTGTG GAGTAATAAC AATCCAGAAG TATCATAGTC 420

55

EP 0 786 519 A2

ACGTCCTCTA TCATCAGATA CACTTGTkAT CACTTGAGTT GGKTTATGGA AkAAAAATGKA 540  
AAATTTGTCT TCTAGTTCTA TTTTAATACC TTCAACTTCA ATCGTATCTG ATGGCTTCAC 600  
5 TTTTGTCTCT AATTCACTGA CAGTCGTACC ATTCACCTTC ACTTTTCCTT CAGAAATTA 660  
AGTTTCTGCC TTACGT 676

(2) INFORMATION FOR SEQ ID NO: 568:

10 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 454 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
15 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 568:

20 GAAACGGTTC TACCAAAAAA CAGTAAGGCG TAAACCCAAT CATGGTAAGA CAAAAAGTAC 60  
AAATAGCCAT GCCCAAGTTG AACTCGCTGT ACGCCTATTT CTTTCTAAAA AGATAATAAT 120  
AAAAGCCAAAT ACTAAATTA 7GATGAATCC AATGGCTAAA ATAATAGTAA ATAACGTTCC 180  
25 TAAATCGTTT GAAATGTAA ATCGCATAGT CTTTCTCCTT ATAAAGAAAG GCACAAAAAA 240  
ACATTTTGCA CCTTTCACGT CATATTATTT ATTCACAGAT AAAGTTAAAA TTGcATTGAA 300  
TCTCTCTCA TTATTGGGA ATGTTCTTTC TTCTATTCT TTAATAGTAA TATTACTAA 360  
30 TTTTAAATTT GTAGCTTCT CAGAACTTAA AAAAGCATTA ATGTTTTTTT CTAATAACTC 420  
KAAAGTCTCA GCTGAAAAAG TTTTAAGTTT AATT 454

35 (2) INFORMATION FOR SEQ ID NO: 569:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 894 base pairs  
(B) TYPE: nucleic acid  
40 (C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 569:

45 ACGATATTAC CACCCTCTGA AATAACATCA ATCATGTGAT CTGTAAAGTC CCAAGGATGA 60  
CTTGTTGTGA AACGAACTCT TGAATCGCT ATTTTAGAAA TTGCTTGTA AAGATCTCCT 120  
50 AAGTCATAIT CTATATCCTG TAAATCTTTA CCATAAGAAT TTACATTTG ACCTAAAAGC 180  
GTTATTCTTT TGTAACCTTC ACGAGCAAGT TCACGTACTT CATCTATAAT GTCTTCAGGT 240  
CTACGGCTTC GTTCTTTACC TCTTGTAAT GGAACAATAC AATATGTACA AAACCTTATCA 300

55

EP 0 786 519 A2

|            |            |             |            |             |            |     |
|------------|------------|-------------|------------|-------------|------------|-----|
| TCAATAACGT | CTCCTCTTT  | AGACCATACT  | TCAACAACCA | TTGCTTTAGA  | TAAGTATGCT | 420 |
| TCTTCTAAAA | TTTCTGGTAA | ATGATGAATA  | TTATGTGTAC | CAAAATATCAT | ATCTACATT  | 480 |
| TGATACGATT | TTAAAAATTT | ATTCACCTACT | GACTCTTCTT | GTGACATACA  | ACCACAAAAC | 540 |
| CGATTAAAA  | TATCAGGTGG | TTCTTTTTTC  | AAATCTTCA  | AATTACCTAT  | TTCATAAAC  | 600 |
| ACTTTGTTCT | CGGCATTTTC | TCTAATCGCA  | CATGTATTAA | TTAAAAAATC  | ATCTGCAGTG | 660 |
| TTAATATCAG | TGCTKGCTTG | aTAGCCTAAT  | GCYTCmAGTA | TACCAGCAAT  | GACCTCAGTG | 720 |
| TCATGTGCAT | TCATTGACA  | TCATATGTT   | TTAATTAAAA | ATGTACGCTC  | GTTCGCCATA | 780 |
| CGCGCATATT | TTYCATCAAT | TtGGgAAATC  | nCTATTATA  | CGAACTTCTk  | GTtTACCnCC | 840 |
| TTTTTTnCGC | TCCTTTAAAA | TTAAGGCGGC  | TGATAAACAG | GTCCAAAAATA | TTAC       | 894 |

(2) INFORMATION FOR SEQ ID NO: 570:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 441 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 570:

|            |            |            |            |            |             |     |
|------------|------------|------------|------------|------------|-------------|-----|
| TATCAATCCC | ACAGCACATG | CTGAACAAGA | TCAAACATGG | GAGAAGATTA | AAGAACGCGG  | 60  |
| TGAACCTAGA | GTGGGTCTTT | CTGCAGATTA | TGCACCAATG | GAATTTGAGC | ATACAGTTAA  | 120 |
| TGGTAAGACT | GAGTAGTCAG | GTGTAGATAT | TGATTTAGTG | AAAAAAATTG | CGAAAGATAA  | 180 |
| TAATTTAAAA | TTAAAAATCG | TCAATATGTC | ATTGTAGTAG | TTGTAGGAG  | CTCTTAAAAAC | 240 |
| TGGAATAAAT | GATATTATTA | TTTCCGAAT  | GACTTCAACG | CCTGAACGTA | AGAAGCAAGT  | 300 |
| TGATTTTICA | GATTCATATA | TGATGACTAA | AAATATCATG | CTTGTAAGAA | AAGATAAAGT  | 360 |
| TAATGAATAT | AAAGATATcm | AAGACTTTAA | TAATAAAanA | GTnGGGGCAC | AAAGGGGACTG | 420 |
| AACCAGAAAA | AATCGCTCAA | C          |            |            |             | 441 |

(2) INFORMATION FOR SEQ ID NO: 571:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1205 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 571:

# EP 0 786 519 A2

|    |  |      |
|----|--|------|
|    | AAGTCGTTCA TCITTTTCTA CAAGGGTGTA AAAATAATCA ATCATATCGT ATAACGCTTC  | 120  |
|    | TTTACCAATG ATGTCATATG GTGTTGTTGT CATTTAATCA CCCATTITCA AAAATTTACT  | 180  |
| 5  | GTTCAGAACT TAAGTTAATA TATAACTAAT ATAACATGAT TTAAACATT TGAAGAGAAAT  | 240  |
|    | ATGCATATTT GCCAATTTAA TTTATATTGT TTGAAAGTGT TTCITTTTTT TTGAAAAAAC  | 300  |
| 10 | GTGGAACCTT ATTTAAAGGT LGATGATGTT CGAGGGTTAG TTCGTTTAAAT AAAGATTGGA | 360  |
|    | ACTTTTGTA ACCTTGATTA TAGTCTTTAA CTTGGAAGTC TAACTCATAA TCCGTAGTAT   | 420  |
|    | CGAAATACTC ACTTTTATCT AAAACCAAGT AATCACCTTT ATATTTAGTT TCTGGGCGAT  | 480  |
| 15 | ATGTCGTTAA TGCACCAAGT AITGATAAAG TTGTATCTTT TACACCAAAAC TGTTCACATA | 540  |
|    | TAATTTGACG AATGTCATCT GGAAGATTGT CGTTTGAAGT AATCAAGTTC ATCTCTGGTT  | 600  |
|    | TAATGTCGAC GATATAGTTG TATTCTAATA GACCAAGCTT TGCTGGTGTC TTTAAAGTCA  | 660  |
| 20 | TTTCATATTG ATTGCTTTTA ACTCTTATGC GTAGTGCAGA GCGATGTTCC TTTAATTGTA  | 720  |
|    | AATCGGGTGT ATCAATATAG TAATTGACTT GCTTAAAGG CACACTGTCT TTAAATATT    | 780  |
|    | TCTCTTGCAA TTTATTATAG ATTGATGCAG TTATCATTTG TTTAAATTCT ATTCATGAT   | 840  |
| 25 | TTGTTGCCAT GATATGTATA CACCTCGTAT CAAATTCAT TATCTTAAAC TATATTATGA   | 900  |
|    | ATGACAAAGT TGAATTTTAA AAGTAATTTT CTTTATCTAT TATCAATGTT AATTGACCA   | 960  |
| 30 | TTAAAAATAG TGTTGTAAG TGTTTGTAT TATTGAATTG TGTAAAAATG TTATGGAATA    | 1020 |
|    | AGAGGAGGAT TAAGCATGAG TTTTATATT AATGAAATTA AAATTAAGA TGACATACTT    | 1080 |
|    | TATTGTTATA CAGAAGATTC TATTAAAGGA TTATCTGAAG TAGGACAAAT GCTCGTTGAT  | 1140 |
| 35 | AGTGATAATT ATGCTTTTGC GTATACATTA GATGATGGTA AAGCGTATGC TTATCTCATT  | 1200 |
|    | TTCGT  | 1205 |

(2) INFORMATION FOR SEQ ID NO: 572:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 570 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 572:

|    |   |     |
|----|---|-----|
| 50 | TGAAGAAATA GCAATGATGA AATGCCGTAT GTTATATGAG ACGGGTCAT TTCTTGAATT  | 60  |
|    | AAGAGAAGAA ACAATTGCTT TATTGAAAAC TGGCATACAA CAATATGATG CATTGATGAT | 120 |
| 55 | TTATTACGTG AAAAGTTTGA TTGGTTTGGG ACAATATTTT GAAGCGGTAG AAGTAATTCA | 180 |

5 ATTTGCTAAG TCAAAATTAA TTGAAGATGA AAAACGATTG ACTCAGTCAT TAGCTGATT 300  
 TGTACGTGA TCAATGAGGG AACAGACGCA CTGATTTTG AAGTTAATAG ACAATGGTCA 360  
 TTTTCAATTT CAAGAAACGG TATTATATAT ATkAAaAyCT AATaCGTACa GTaTATAACCT 420  
 CATTAGTTTA ATGATTGAGT ATTTAAGGTT CGCAAATTGT ACACAAGAAC TGACAATTGA 480  
 10 AAAGTATGGT ATGGATGTAA CTTTGTGACC AGCTAATTTA AAAGGGCTAG AACATACAAC 540  
 ACTTAAAGAA AAGTTATATC CTAACGTTAT 570

## (2) INFORMATION FOR SEQ ID NO: 573:

15 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 939 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

20

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 573:

25 GTTGAATGGT TAGCAGCTGC AGTTGTATTA TATTCTGTG GTGTAATGT TGACGCTCAT 60  
 GTATCATTCa TGTCCTTTAT TGCAATATTT ATCATTGCTG CATTATCAGG TTTAGTCAGC 120  
 TTTATTCTCG GTGGTTTCGG CGCTTTCGAT TTAGTTGTAT TACTAGGATT TAAAACTTTA 180  
 30 GGTGTCCTTG AGSAAAAAGT ATTATTAATG CTACTTCTAT ATCGTTTTCG GTACTATTTT 240  
 GTACCGGTAA TTATTGCATT AATTTTATCA TCATTGGAAT TTGGTACATC AGCTAAGAAG 300  
 TACATTGAGG GATCTAAATA CTTTATTCTT GCTAAAGATG TTACGTCAIT TTTAATGTCT 360  
 35 TATCAAAAGG ATATTATTGC TAAAAATCCA TCATTATCAT TAGCAATTTT AGTATTCTTT 420  
 ACAAGTATGA TCTTTTTTGT AAATAACTTA ACGATTGTkt ACGATGCTTT tATATGATGG 480  
 AAATCACTTA ACGTATTATA TTCTATTGGC AATTCACTACT AGTGCTTGTT TATTACTTTT 540  
 40 ACTGAATGTA GTTGGTATTT ATAAGCAAAG TAGACGTGCC ATTATCTTTG CTATGATTTC 600  
 AATTTTATTA ATCAGATGG CGACATTCTT CACTTACGCT TCATATATTT TAATAACATG 660  
 45 GTTAGCTATT ATTTTGTTC TGCTTATTGT AGCTTTCCTG AGAGCGAATA GGTGAAAACG 720  
 CCCAGTAAAG ATGAGAAATA TAGTTGCAAT GCTTTTATTC AGTTTATTTA TTTTATATGT 780  
 TAACCATATA TTTATTGCTG GAACGTTATA TGCAATTAGAT ATTTATACGA TTGAAATGCA 840  
 50 TACATCTGTA TTGCGCTATT ACTTCTGGCT TACGATTTTA ATCATCGCTA TCATCATAGG 900  
 TATGATTGCA TGGTTGTTTG ATTATCAATT TAGCAAAGT 939

55

## (2) INFORMATION FOR SEQ ID NO: 574:



(A) LENGTH: 1059 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 574:

|    |  |      |
|----|--|------|
| 10 | GAATTAATTA AATATTACAC ACAGCCATCAT TTITCATTTT CAAATAAATG GCTATATCAA | 60   |
|    | TATGATAATG GAAACATTTA TGTGAACTT AnGAGATATT CATGGTCAGC ACATATATCT   | 120  |
|    | TTATGGGGCG CTGAAaGTYG GGGAAATATT AATCAGTTAA AAGATCGTTA CGTAGATGTG  | 180  |
| 15 | TTTGAGCTAA AAGACAAAGA TACTGATCAG TTATGGTGGT CTTATAGAGA GACATTTACA  | 240  |
|    | GGTGGCGTTA CACCAGCCGC AAAACCTTCT GATAAACTT ATAATCTTTT TGTGCAATAC   | 300  |
|    | AAAGATAAAC TACAAACGAT TATTGGTGGC CATAAAATAT ACCAAGGCAA TAAACCACTA  | 360  |
| 20 | TAAACATTGA AAGAAATCGA TTCCCGTGCA CGAGAAGCGT TAATAAAAAA TAAATATTA   | 420  |
|    | TATAACGAAA ATCGTAATAA AGGTAAGCTT AAGATCACCG GTGGCGGTAA TAACTACACT  | 480  |
|    | ATIGATTTAA GCAAAAGATT ACATTCAGAT CTAGCAAATG TTTATGTTAA AAATCCTAAT  | 540  |
| 25 | AAAATAACTG TTGACGTCCT CTTTGATTAG TATAATGAAG TGACTTATAC TTCATGCACT  | 600  |
|    | TTAATTCCAA ATCAGATTAT TTAATGATA ATTTTAAAG TGTATGATGT ATATAATAGG    | 660  |
| 30 | TAAATTTTC TATATATTTA AATGGAAATG GGAGTAGGAA TGTGACAGAA ATAGTATTTT   | 720  |
|    | ATAAAATTTA TTCtTGTCAC TCCCAACTT GCACATTATT GTAAGCTGAC TTTCGCCAG    | 780  |
|    | CTTCTATGTT GGGGCCCCG CAACTTGCAT TGTCTGTAGA aTTTCTTTTT GAAATTCCT    | 840  |
| 35 | ATGTTGGGGC CCCGCTTATA ATTGAAAAAT GCTTGTTACA TGGGCATTIT CATTCGGTCA  | 900  |
|    | ACTACTACCA ATATAATATT GtAGaGCCTA AGACATTGAT TTATATGTCT TTAGGCTCTA  | 960  |
| 40 | TTCCCTCAIT TAATGATTAA nTTATTATAG CAATACTTTA TTGTCCCATG ATTAGTGTTC  | 1020 |
|    | TTTTAATGAG ACATAGTAAC TATAAAGTTT AATAATCGT                         | 1059 |

(2) INFORMATION FOR SEQ ID NO: 575:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 574 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 575:

|    |  |    |
|----|--|----|
| 55 | GTTCGCTTGA GGTCTGTTT CATATTATA CTTCGAAGGA TTTACCTTTT TgAGTCTGG | 60 |
|----|--|----|

EP 0 786 519 A2

|    |  |     |
|----|--|-----|
|    | TTCAACTKGC TTTTATTCT TTTCGAAATC AGCTGGTTGA GTAGTTATGA GTTCATTATT   | 180 |
|    | TTTATTAGmA TAAATCTTAC CATTAAACATA TTTATAATCT TTTGTTATAA AGTCACCATT | 240 |
| 5  | TCTGAATGGA ACTACTTGAT TATGACCTTT AGAGAATAAA TCAGTACCGA ACATTAAATA  | 300 |
|    | GTCTCTGGTA TCTATACCAG CCAATGTAA AATTGTTGGC ATTACATCGA CTTGACCAGC   | 360 |
|    | ATATTCAATTA TTGATACCAG CAGATTTACC AGGGATTTTA ATCCAGAAAC CAGTTCTGTT | 420 |
| 10 | TAAATCTGTA AATTTAGCCG GTGTGATTTT TTCACCTAAT AGTTTTTCCA TGGCATTGTT  | 480 |
|    | ATGGTTTTTCA GAGATACCAT AGTGGTCACC ATAAATCATA ATCACTGAAT TGTCATATAA | 540 |
| 15 | TCCTTTTTTC TTCAAGTCAT TAATATATTC TTCT                              | 574 |

(2) INFORMATION FOR SEQ ID NO: 576:

|    |                               |
|----|-------------------------------|
|    | (i) SEQUENCE CHARACTERISTICS: |
| 20 | (A) LENGTH: 796 base pairs    |
|    | (B) TYPE: nucleic acid        |
|    | (C) STRANDEDNESS: double      |
|    | (D) TOPOLOGY: linear          |

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 576:

|    |  |     |
|----|--|-----|
|    | CAATGTTTTA TAGTACAATA TATTTnAAT AATACTCGTT AAGGAGAAAT ATATGATATC   | 60  |
| 30 | AATTCAACGA ATTTCGACAG GAAAAATCCA AGATTTCCT TATAGCTCGA AAAGACCGAT   | 120 |
|    | GCGCTCTGCT TTAGATAAAA CTAAGATTTT ACAAAACAACA TGGTTATCTT CAACTGGTTT | 180 |
|    | CACTGGTGAT GAACAGGCTT ATAAAGATCA TGGTGGACCA CATAAGACG TTTGTGGGTT   | 240 |
| 35 | TAGTAAGCAT AATTATGCAC TGTATCAAGA TGATTTACCT ACACACCTA CTCATCGCAT   | 300 |
|    | GTTTGGAGAG AATTTAACAT TTGATTATTT AGACGAATCT GATGTTTACT TTGGTAATCA  | 360 |
|    | ATATCGTTTA GGTGAAGCGT TAATTGAGGT TTCTGAAAT AGAGAACCAT ACTGAAAAAT   | 420 |
| 40 | TCAAGCAAAA TATAATATTC CTGATTTAGT GAAGCGCATG TCTACATCTG GTAAAAACAG  | 480 |
|    | TTTCTATTTT CCGGTATTAA AACAAAGGCTA TGTATCTCCA AATGATCAGC TTTACTTAAT | 540 |
|    | ACAAGAAGCA CCAATCGAAC ATCGTTTATC TGTACAACAG CTTAATGACC TTTATTATAA  | 600 |
| 45 | TGATAGACAA AATCAAGaTA TGTACGATA TGCATAAAC AATCCATTTT TGTCACCAAC    | 660 |
|    | AAGACGCGAT AAACCTTCAA AAATGTATAA CAGAACATTG GAAATAATTA CCTTTCATTn  | 720 |
| 50 | ATAAGTGTTA AATGAACCTT TCAAAACAnA AAGGAATCAA CTTACACAT CGTTTGTATG   | 780 |
|    | AATAGTCTTA TCTATA  | 796 |

(2) INFORMATION FOR SEQ ID NO: 577:

55

- (A) LENGTH: 1095 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 577:

```

10 AGAAATTATG ACAAATATA AAGATGGAAA GTTAGTTTAT GCATCAGTCG AACCAGGATC      60
   TTAOGTAATA CATAAGATG ATGCAATTAA ATATGACGAT TATTCTAAGT TAAAAAAATT      120
   AAGTCAGCTA ACTAACTTG ATCATCCAAA ACCAGTTCCA TATAGCGTAc TCAAAATCAAA      180
15 TCTTTGGAG TACCTTTAAC AAGCGTTTCA TTTATGACAC ATGGATCAAA GGATACTAAA      240
   GATGAAGTGT TGCGGCATT GGCCTATTTC ACTTTTTCAC CAAAAAATTA TGAAGACAAG      300
   TCTAAATCCAG ATCCAAAAGT TTTAAATTIA GTACATATGG ATTTCTTAAA TGCATCTAGT      360
20 GATTTTGGTA ACGCACATTT TGTGTGTTTA AGTAAATATA TTAAGAGTA TGAATCAAAC      420
   TATGAACAG CGTCAGATGA TTCCTTAAAA TAGTATTTAC TGTGTGAAAA ATAAATAGTG      480
   TACTACATTA AATAATCGCA ATAATAATCC CGATAAACAA TCAGCATTAC TGCTTATCAC      540
25 ATAGAGTTCG TAATAACTAT AACTCTATGA TTCGCAAATA ATAAATGATT GTCATCGGGA      600
   TTTATTTTTA TCAATTATTA AAGTGACATT ACCTTGTTCA TCAGCAGGTT TGAAAAACAGT      660
30 AATCACTGCA CTAATAATTG CTAATAATGT TGGATACCT GTCCAACAGA ATATTAGTG      720
   TAGAATACCT TGCATATTCT TGCGGCATA AAATTTATGA ATACCAAAAC TACCTAAGAA      780
   CAATGCTAAT AAAATATAAA TAACTTTGTT TACTTGCAAT TCTTCCCTC CAGTTGAATT      840
35 GCTTATAATG ACATTAGCTT CTCITTTTAT TATACCCACT TTAGTTCAA ACATTCTAGT      900
   TTAAGCATTC CCAATCATCT AAATTCAGT TATTCAATCC TTACAATAAA TTTAGGATTA      960
   CATTTAGATT GCATTGTATT ATTTTACGTG TGAAATATAC GTAATGAATC ACATGACAay    1020
40 CTyCAATTG AAAAATATAC ATTTATGAT GTAAGGTCG ATTTTAAATA TATTTACGTh    1080
   AAAATAGTTT GGATG                                     1095

```

## (2) INFORMATION FOR SEQ ID NO: 578:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 489 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 578:

ACTGCCAAAA TAATCATAGC CATGTTCTAC AGCTGCTTTC GCTACAATAT CCAAACGCAT 120  
 TTCAAAAACAA GCGGTACAAC GTAAGCCGCC TTCTTTTTCa TCAGCTAATT CTTTATCCTT 180  
 5 CACCATTTCt ACAAACCTAT GTGGTTCATA AGGTGCTTCA ATATACTTCA CATTGCGACC 240  
 AGTCTTGCGA TTAANAATCTT CCACAAATTG TTCTTGCACT TTAGCACGTC GTAAGTACTC 300  
 10 ATTTTTCGGA TGAATATTTG AATTCGCGAA ATAAATTGCA ATGTCTGCAT ATTGTGTTAA 360  
 AAACCTCTAAT GTATATGTAC TACAAGTGTC ACAAACATA TGCAATAAGa TTTTAGGTCT 420  
 GATGCTCTCT CTTTCCCACT GsCGATTAA TnTCTTCAAC ACCTTGTCa TAATTAAITT 480  
 15 GTnGATTIn 489

(2) INFORMATION FOR SEQ ID NO: 579:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1287 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 579:

CGTACTGCAT CATTGTGTA AAAATTGCA AACTTCTTCA AGATGTTATG GCTTAGACTT 60  
 30 AAAGCGATGA AGCACTACAA AGCCTTAAAT AAAGAATCTA AGAAGCAAGA ATTTGAAAAT 120  
 TCATTCAAAG ATGTTCAAAA AATTATGCGT ATTGTGAATC ACAATATTAT TTTACGCTTA 180  
 AAAGAAGAAC AAAATAGTAC AAATGTACTT GAGGTTAGCT TAGTCATTAA TCATTACTAT 240  
 35 GATATGAGTC GCTCATTAAA GTGGCGTGCA CAACGTCGAA AAGAAGCTCA AGAAAAACAGC 300  
 AATCAAAATCA TACGCAAGC TATGTTCCAT AACCAAAAT TGAAGCATT GTACTTACAA 360  
 CGTCATCTTT TAGATGAATT AATTGCGAAA AATAAAATCA ACAATATCGT TGCAGCTCAA 420  
 40 ATTCGAGAAA ATATCAATTa CAACGAAATT GTCTTGTCCT TACAGTCCAA ACATTAGCA 480  
 AGaCwTaCaw TmCCCCGTA CATATGAGAC AAAGTCATTa TCATCTCATA TGTACGGGT 540  
 TTTTATATTC AACATCAAAA AATCAGATTG ATGAAAAGTA AATAACCTTT CATCAATCCG 600  
 45 ATTTGATTAT AGAATCTATT TTTTAAGTTT AAATGGAATT GTACATACGT TAACATTCCT 660  
 TTGATAAATT AAATACAATT TCATACGCAA ACTAGTTTGA TTGTGAATA AATTATGCCA 720  
 50 ACGTTTCTTA GTAATAAATT CTGGTATCAC TACTGTAATC ATATAGTTTT GATCGTTGGC 780  
 TTTACGATTA ATCTTATCGA TAAACGTGA AATTGGTCGT ATAATACTGC GATATCCGA 840  
 ATGTAANAATA ACTAATCTTA CATCTGGGAA ATGACGTTTC CATTTCTCTT GGAATGCTTT 900

EP 0 786 519 A2

|    |            |            |            |            |            |            |      |
|----|------------|------------|------------|------------|------------|------------|------|
|    | TGCGTAATAA | ATAGACTTAT | CAACTGCTGT | TGTAATACCT | GTTATCGGCA | CAATTGCTAA | 1020 |
| 5  | ATTAGATCT  | ACCACGTCCA | CATTAAAGAC | ATCAATGTCA | GAACGTAACT | GTTCTGCGAT | 1080 |
|    | ATCTOGATAA | TGKTTGKATA | TTTTCAAGAA | AAAAATCACC | ACGAACGGCA | AGAAAAATAG | 1140 |
|    | TATCGGCCAT | ACTTGGCTAA | ATTTAGTTAT | GAGTAAAACT | CATAANAACA | ATAAATGTCA | 1200 |
| 10 | CGATACACC  | AAGTAAGTTC | CAACGCAACT | TACTTAACCA | ATTCTTAGGA | CGTTCATGAA | 1260 |
|    | TCCATTTAAT | AAGTACACG  | AATTGTG    |            |            |            | 1287 |

(2) INFORMATION FOR SEQ ID NO: 580:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1223 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEO ID NO: 580:

|    |            |            |            |             |            |            |      |
|----|------------|------------|------------|-------------|------------|------------|------|
| 25 | GCTCTCTTC  | CTATATCCAT | GGGTGTCTTA | GTTGGGAATG  | GCTTATTCCT | TATCCCCAAC | 60   |
|    | ATCTGTCAAT | TCAAGACTAT | CAGTGTATAC | TAAATCGCCT  | hTCTATACnT | CCATnCATT  | 120  |
|    | CATTTTCCAG | TGGGAACAAT | TGTTTAAAT  | ATTTTAAATA  | TCATTATCAA | TCCGAAGATA | 180  |
| 30 | TTTACTATT  | ACTTTTACAA | TAATCAATCA | TTTAACTACA  | CATGGGTAGT | CGGTGGTTTG | 240  |
|    | CTGGGGTGA  | GCITTTTAAC | TGGCAATTTA | TGTGTATTGC  | CAAAATTAGG | TGCAACATTA | 300  |
|    | ACTGTAATTG | CAACAGITGC | GGGTCAAATT | ATTATGGGTG  | TCATTATTGA | TACATTGGGA | 360  |
| 35 | TTATTGGCG  | CTACAATTCA | TGATTTTAAT | TTAATTAAAG  | CAATTGGAGT | ATTGTTACTC | 420  |
|    | ATTGTCGGCA | TGTCATAAT  | GAATCAATTT | AAACAAGAATA | ATTATTACT  | AACTGATCAA | 480  |
|    | AAGTATTTAC | TGTTTTGGCT | TCTATTAGGA | TTTATTTTGT  | GTTTCTTCC  | ACCTATTCAA | 540  |
| 40 | ACGACAATTA | ATAGTGCTTT | AGCTAGTCAT | ACTCATTAC   | CAGCCTTTGC | ATCATTAGTA | 600  |
|    | TCATTTACAA | TGGGTCAAT  | AGCGCTATTG | ATTTTAAACG  | CTATTTTAA  | TCGTTCTTTA | 660  |
|    | AAACTAAAAA | CAAGTCATT  | AAAAITCGGT | AAATTAAAGC  | CTATCTATT  | TACTGGCGGT | 720  |
| 45 | ATACTTGGCA | TGGCTTTTGT | AACAGCTAAC | ATTATCTTAA  | TGCTCATAT  | GGGTGCAGCA | 780  |
|    | TTAAACAAC  | TTATTGGCAT | GTTTGGCCAG | ATTCTAATGT  | GCATATTGAT | AGATCACTTT | 840  |
|    | GGATTATTTG | GTTCACTTAA | ATAGCAATG  | ACATCCAGAA  | AAACTATTGT | TCTATTATGT | 900  |
| 50 | ATTTTGACAG | GCATTATCAT | TTTAAAGATA | TTTTAAATTA  | ACTTTTAGCT | TATCATTTTA | 960  |
|    | ACITGTAATT | ATTTTAAAAA | GTGATAAGCT | ATTTTTTGTG  | GGTCTAAAAA | TCITTAGAAA | 1020 |

CAACTCATT C TTAAGACCTA AATTAATGTT ATnTTTTAAT AATTACACC AAATTAATAG 1140  
 CAAAAATTAT GTTATTGGTG CTAATATTTC ATAGTTGGTT ATTCAATTAA TTAATAATAA 1200  
 5 GTCAAAATGC ACAACTTTTT ATn 1223

## (2) INFORMATION FOR SEQ ID NO: 581:

(i) SEQUENCE CHARACTERISTICS:  
 10 (A) LENGTH: 454 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 581:

CACCTTTTTA TCATGCTTAG TTATCAATAA ATCTATATTT TGCTGTTTTA CAATTTTTTT 60  
 20 AACTTTATCA ATCTCATTAT CTGGACTAA ATAAATATAT GATCTTGCAT CTGTGCTAG 120  
 AGCTTGTTGG TGTTTTCTG ATAAACATA TGTGATGAA GGTGAATAA TAATGCCTAA 180  
 TGTACAAAAA CTGATAATTA ATATACTGCT TATCAATAAC ATTAAGCGGT GGTGAAACTT 240  
 25 CATCATGTT CTTTAGTCT TTCCAATTTA TAGCCTAAGC CACGCACAGT TTTAATAAGT 300  
 TGTGCTTCT TAGGATTATC TTCTAATTTA TCTCTAAAT GACTGATATG TACATCAACA 360  
 30 ATTCCTGAGT CTCCTGCAA TICATAATTC CATACCGTAT TTAACATATG CTCTCTCGTA 420  
 ATGACTCTGC CTGTCTTTT TATCAATAA AGCA 454

## (2) INFORMATION FOR SEQ ID NO: 582:

(i) SEQUENCE CHARACTERISTICS:  
 35 (A) LENGTH: 452 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 582:

TTTATAAGAT TTTATTTACA AATAATTGGT TTTCATATGT ATAAACACTT TTGACTTTCA 60  
 45 AATCTTAAGC ATAATTCTAT TACAATACAA TCCCTTATTA GAATGATTTA TGTAATAAAA 120  
 AAAGCGGAGT TTCCCCGAGC TTTTCTAAAC GACTACATAA AATATAAGAT TGCAATTAAA 180  
 50 TGCAATAGTG ATGCTATTAC AATAAAAAA TGCCAAATCA TATGAAAATA TGGTCTATTTC 240  
 TTTTGTGCAT AAAACCATGC ACCAATTGTA TAAGACACAC CACCTAAGAA AATGAATAAT 300  
 ATGAATATCC ATGATGTGCG AATAAAAAA ATTGGTAACA AGATAATACC TACCAGCCCC 360

55

AAATCCCCC AAGTGTGCTGTC TCACATAAT AA 452

(2) INFORMATION FOR SEQ ID NO: 583:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1472 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 583:

|             |            |            |            |             |            |       |
|-------------|------------|------------|------------|-------------|------------|-------|
| CGCTTTTTCG  | AATAAGTCAT | TAGCCGCTTT | TAGCGCTTCT | TCTTTTCGAT  | CTACAACAAG | 600   |
| TAAAAATAAT  | GGCTTTAAGC | CTTCTTCTTT | TTCACTTTCA | AGCATATCTG  | GTTTTTGAAC | 1200  |
| CATTTCAAAT  | GGAGATTACA | ATCCATTATT | ATCGCTCATT | TCAATTAATG  | CATCATCTG  | 1800  |
| TGCTTGTGAC  | ATACTTGCAC | TAGCCTGTTT | TGCATTTTCT | TGAAGGAAAT  | ATAAGTTTTC | 2400  |
| CAATTTAGGA  | TGCTTTATTA | ATGTACTTAA | TGTAATCGGT | GTAATGTCTT  | TCTCATAGA  | 3000  |
| CACCTCAACT  | ACTGTACTAT | TTGTTTCACT | AGGAATTGGT | GTTTTCATCT  | GAATATGCTT | 3600  |
| TGATACTTCT  | CCAATTCCAA | CGACAGATTG | ATTTTTCGTT | CGATTATAAA  | AAATAATATT | 4200  |
| GTGCGCTTCT  | TCTAAGTAG  | TATAAAAAAT | ATAACCATTA | CGTTTAATAC  | CGTTGTACGT | 4800  |
| GTGCGTATAA  | ATCGTATATT | GGTTTCCAGG | TTCAAATTCT | TCAGTTTCAG  | CTAAAAAGAA | 5400  |
| ATAACGCGGT  | ATCTTAAATT | CGCCTTACAC | AAGACCACTT | ATTAATCAAA  | ACTCTTCTGC | 6000  |
| AGTGATTGGA  | TTGAACAATG | TCTCTTTCAT | ATTACTTATA | CGAAATTCAC  | AAGCTTCACT | 6600  |
| ACGCTTTAAA  | TAACTCTGCT | TTAATGGTTT | CAATTGTGTA | TTAAAACGAA  | ACTGTACACG | 7200  |
| TATTTTATTT  | TGTGCACCTG | TTTCAACACT | AATAATTCTA | CCACATCCAA  | GTAGTCCAGT | 7800  |
| ATCGGTCTGA  | ACTTGATAAA | AGATGACTTG | ATCTCTACTA | TTAGCGTTTT  | TAAACGCTCT | 8400  |
| AAATCCTTGA  | GATGGGTTAA | AATGTGCGCC | TGATTCAAAT | AAAGCTGTTT  | GTCTACTATA | 9000  |
| CGGTTCAATTA | TGATTCCAAC | GGTTATATCC | ACAATTCAAC | CAAAAAATAA  | TCGTTTCTGC | 9600  |
| TGTCATCTTA  | ATACTCCTTA | ACCTGAATAA | ATTTTAGAAA | CACATTAAGT  | TACATTCTTT | 10200 |
| TAGTGTTCCT  | TATGCAAGTG | GACGCGIATG | CGAACCAACT | TATACCCCTT  | GTTCACTGGG | 10800 |
| ATTTTAATCG  | CATTTTCTAT | AACATTGTAG | CGCCACGAGC | ATTAATTTCG  | GTCCACAGAC | 11400 |
| CTTATCGTTT  | TCATTCTCTA | TGAGTCGAAA | CTATTITGCT | TTACAACAAG  | TGCGAATCTA | 12000 |
| ATAACAGTTG  | GACACACATA | CGAGCAACTG | TATACCTTTT | AATCAGTTTT  | CTATATTTTA | 12600 |
| TTTATTATAT  | CTGTCTTAAT | GATAAAAAAT | GTTACAAACA | GTTTAAACATA | TTTAGTCACT | 13200 |





TAAAAAGAAG ACTATTTTT ATTACAAGAA AATGTATCTA GTAAACTTAA AGTAGCAAGA 180  
 CCTAATAAAT TTAATGCATG TTGTGCACCT TTTTACCTT GCCAGCCTT GAAATGTTTG 240  
 5 TAAGCAGCTA CACTTAAAT GCCTATCGTT GATAGTGATG CAAGGCGAGA AATGTTTTTA 300  
 TTGATAAAGC TAGCTGAGTA TAAAGCAGCA GTAGTTGCTT CTGCAATGCC GACGTATTTT 360  
 10 ACAAGTCTTT TTTGCAAGCC AAAAGTATGT TCAAACAGTT CAATCATACC CTTATCTTCT 420  
 TGCAATTAGG GTTTACTGGC TTGTATAGC TCTTTCGCAA GTTTTAAATT CGTTGCGTAA 480  
 CGCAAAATCA TATTTAATTC CTCCAATAT TTGATTTTT GTGAAGATG ATTACTTTAT 540  
 15 CTTTTTACC CGTTTCTATA AAAATGAATC AATTATGTAA CGTATGTGA GTTTAGGAAT 600  
 GTTTGCTATG GAAATATAAT TCTGTTCACT CAAATGTAT GAAATTAAAT TGATGTTTTG 660  
 TCGAGTTGCT CTTTTAATTT GGTTAGATTG TTTTITAGAG AAGCGGTACT ATTTTAAAGT 720  
 20 GCATCAACAG ATTTACCTTC GTTTTGAGAC ATTGAGTTTA TTACAGCACG AAGTTCTGTT 780  
 TCTAGTATGT CagCGTCGCT TTAGCATTAG AACTTAaTat TTaTAcTCTT 830

## (2) INFORMATION FOR SEQ ID NO: 586:

25 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 412 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 30 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 586:

35 TTAGGAGCTT TTTACAATCA GTACAATGAG CTCATTGTTA TTAGTCCTTT AACGCGTCT 60  
 TTTAATGCTG GCGCTACATT TGGCGAATT CATCATTTAA TTGATACTGA AACTTTAGCA 120  
 AAATTAGAAC ATGAAAAAGG ACATTATTAT CAGAAGATGA TATGTGATGA CAATGTAGAA 180  
 40 ATGATTTCTA TAAATAACAT ACCGAAATAT CCGAGAAATC ATAATGTATT AACTAATCAT 240  
 GACTCATACG AATATTCAAT GAATTTAGGA AGTAGTAATA GTTATTCAAA GTATGAGCTT 300  
 ACCTTAGATG ATATTTATGT TGGTGCTACC TTTTAAACAA TTATATTTAT ATTCTAGCCm 360  
 45 ACTAAATAAA AGGGKaCtaT TTGaATCaAA CmAtaTGtAT TAACCTTTTT TA 412

## (2) INFORMATION FOR SEQ ID NO: 587:

50 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 4709 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 55 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 587:

|    |  |      |
|----|--|------|
|    | TTCAAGTTTGA AATTAATCAT ATAAATTTCT TATGGGAGGG TTGATATCTT AATGATTAAC | 60   |
| 5  | ATTATTTTCAG CTATAGGATC TATTGGAACA TTTATTATGG CTTTATTTTA TTTTGTATCA | 120  |
|    | GTTTCAGTTC AACCTTTATCA AATGAAATTT AGCTTTCTGC CAGCTTTAGG TTTTAAACCA | 180  |
|    | ATTTTATTAG AAAGGGAGGA GGATCAACTT AATATAATGA ATTCGGCAAC AGAAGAGCAT  | 240  |
| 10 | CATCATAAAG ATTATATTAA ACTATATAAT TTAGGTGGCG GTGCTGCTAA TAAAATTGCA  | 300  |
|    | ATAGAGGTTT TATTGGGGAA GGATATAGTC ATTGAGAAAA AATACGTGCA TATTGTACCT  | 360  |
| 15 | AGTAAGAAGG GGTACATGTT ACCAATTAAT AAAAATGTGT ACGAAGAATT AGAAAGAACG  | 420  |
|    | ATTGAsAACA ATGGTCATGA AGCTGATTGG AATGTACGTA TGACTTATTA TCATAATGTA  | 480  |
|    | AGTCGCAAC AACAGGAAGT TATATTAAAA GGTCAAAATG ACCGTTTTAA TACTTATAAT   | 540  |
| 20 | AATAAGAAAA TTTATGATTT GCAGTTTATC TAAAAATGTA TTTAAGAGGG TAGTTGTTTA  | 600  |
|    | TTGGCAAAAA TATCAATCAA TTTTAATGAA ATAATGGCGT CATTACTATA AAATATTACT  | 660  |
|    | TTATGTTGTA ATGCAITTTT CTATAAGATA GAACTAAAG GAGGGGCAAA GATGCAAAAT   | 720  |
| 25 | AGACAAATAC ATCAACATGA CTTTGCTCAA GTGGACCACT TAATTAGAAC GGCATTTGAA  | 780  |
|    | AATAGTGAAC ATGGTTATGG TAATGAATCA GAGCTAGTAG ACCAAATTCG TCTAAGTGAT  | 840  |
| 30 | ACGTATGACA ATACCTTAGA ATTAGTAGCT GTTCTTCAAA ATGAAGTTGT AGGGCACGGT  | 900  |
|    | TTACTAAGTG AAGTTTATCT TGATAACGAG GCACAACGGG AAATTGGATT AGTGTTAGCA  | 960  |
|    | CCTGTATCTG TTGATATTCA TCATCAAAAT AAAGGTATTG GGAAGCGATT GATTCAAGCA  | 1020 |
| 35 | TTAGAACGAG AAGCAATATT AAAAGGATAT AATTTTATCA GTGTATTAGG ATGGCCGACG  | 1080 |
|    | TATTATGCCA ATCTAGGATA TCAACGGCGA AGTATGTACG ACATTATATCC ACCATATGAT | 1140 |
|    | GGTATACCAAG ACGAAGCGTT TTTAATTAAA GAATTTAAAG TGAACAGTTT AGCGGGAAAA | 1200 |
| 40 | ACAGGTACCA TAAATTACAC ATCTGCTTTT GAAAAAATAT GATTTCAGC TAGGATTACA   | 1260 |
|    | TTAGGTAGAG TTCATATTAA TAATAAAAAA TGTITGCAAT CAAATCGTAC GTTGTGTTTT  | 1320 |
|    | GTAATCTCTA AAATAGCAAT AAATAAAATG TTTGTTAGTA AAGTATTATT GTGGATAATA  | 1380 |
| 45 | AAATATCGAT ACAAATTAAT TGCTATAATG CAATTTTAGT GTATAATTCC ATTGACAGAG  | 1440 |
|    | ATTAAATATA TCTTTAAAGG GTATATAGTT AATATAAAAT GACTTTTTAA AAAGAGGGAA  | 1500 |
| 50 | IAAAATGAAT ATGAAGAAAA AAGAAAAACA CGCAATTCGG AAAAAATCGA TTGCGCTGGC  | 1560 |
|    | TTCAAGTCTT GTAGGTACGT TAATCGGTTT TGGACTACTC AGCAGTAAAG AAGCAGATGC  | 1620 |
| 55 | AAGTGAAAAA AGTGTACGC AATCTGATAG CGCAAGTAAC GAAAGCAAAA GTAATGATT    | 1680 |



EP 0 786 519 A2

|    |  |      |
|----|--|------|
|    | CGATTCTGAC TCCGACAGIG ATTCCGACTC AGACAGCGAT TCAGATTCCG ACAGTGATTG  | 3600 |
|    | CGACTCAGAT AGCGATTCCG ACTCAGATAG CGACTCAGAT TCAGACAGCG ATTGAGATTG  | 3660 |
| 5  | AGACAGCGAT TCAGATTGAG ATAGCGATTG AGATTCCGAC AGTGACTCAG ATTCCGACAG  | 3720 |
|    | TGACTCGGAT TCAGATAGCG ATTCAGATTG CGACAGTGAC TCAGATTCCG ACAGTGACTC  | 3780 |
|    | AGACTCAGAC AGTGATTCCG ATTCAGCGAG TGATTCCGAT TCAGATAGTG ATTCCGACTC  | 3840 |
| 10 | CGACAGTGAC TCCGATTGAG ATAGCGACTC AGACTCGGAT AGCGACTCGG ATTCAGATAG  | 3900 |
|    | CGATTCCGAC TCAGATAGCG ATTCAGAAATC AGACAGCGAT TCAGATTGAG ACAGCGACTC | 3960 |
|    | AGACAGTGAC TCAGATTGAG ATAGTGACTC GGATTCCGCG AGTGATTGAG ACTCAGGTAG  | 4020 |
| 15 | TGACTCGGAT TCATCAAGTG ATTCCGACTC AGAAAGTGAT TCAAATAGCG ATTCCGAGTC  | 4080 |
|    | AGTTTCTAAC AATAATGTAG TTCCGCTTAA TTCACCTAAA AATGGTACTA ATGCTTCTAA  | 4140 |
|    | TAAAAATGAG GCTAAAGATA GTAAAGAACC ATTACCAGAT ACAGGTTCTG AAGATGAAGC  | 4200 |
| 20 | AAATACGTCA CTAATTGGG GATTATTAGC ATCAATAGGT TCATTACTAC TTTTCAGAAG   | 4260 |
|    | AAAAAAGAA AATAAAGATA AGAAATAAGT AATAATGATA TTAATTAAT CATATGATTG    | 4320 |
| 25 | ATGAAGTAAC FCTTAAAGG GTGGCTTTT TACTTGGATT TTCCAAATAT ATTGTTTGAA    | 4380 |
|    | TATAATTAAAT AATAATTCA TCAACAGTTA ATTAATTAA AAAGGTAGAT GTTATATAAT   | 4440 |
|    | TTGGCTTGGC GAAAAATAG GGTGTAAGGT AGGTGTGTAA TTAGGAAAAA TTAAGGAGAA   | 4500 |
| 30 | AATACAGTTG AAAAATAAAT TGCTAGTTT ATCAATTGGG GCATTATGTG TATCACAAT    | 4560 |
|    | TTGGGAAAGT AATCGTGCGA GTCCAGTGGT TTCTGGGAG AAGAATCCAT ATGTATCTAG   | 4620 |
| 35 | TCGTGAAAC TGACTAATAA TAAAAATAA TCTAGAACAG TAGAAGAGTA TAAGAAAAA     | 4680 |
|    | TTGGATGATT TCAATATGGT CcTTCCCA                                     | 4709 |

(2) INFORMATION FOR SEQ ID NO: 588:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1554 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 588:

|    |   |     |
|----|---|-----|
| 50 | CITTTTTAAT TAwCgGaAtA TTGtCaTgaT tAcAcTTCGT TAGGGTTTAc gTCGtAAATT | 60  |
|    | CATTTaaTAA gCgCTTCAcC ATTAAtGTG GTAmCCTTTA ATTGcCAGT AGAAACATCA   | 120 |
|    | CAGTAACtAA GOGCAATTC AGGTTGATTC ATAACAAAC TTAaAATATA GTTATTTTGT   | 180 |

55

CGTCTAACCA TACCTTTCGT TTGTTTCGGA TCTCCACTCT GTTCACAAAT AGCTACTTTA 300  
 TATCCATTAT TAACAAGTGT ATCTATATAA CTATCTGCAG AATGATACGG AACACCACAC 360  
 5 ATCCGAATTG GATTTTCTTT TTTAGCATCT CTTTATGTTA AAGTAATTTC AAGTAYACGT 420  
 GATGCTCTCT TGGCATCTTC ATAAACATT TCATAGAAAT CACCTAGTCT AAAAAATAAT 480  
 AAGCAATCTT GGTATTCTGA TTTTATTTT AAATATTGCT GCATCATGG TGTAAACATT 540  
 10 GACATATTAT TTCTTCACAA CCCTGTCTCT TTTTAAAT TTGTCITTAC AATATATTCTG 600  
 TTGTGAAGyT TTTTAATTAT TAATTATTIA ACTTATACAT TTTAACATAC TTACTTTTAC 660  
 AAACCTATTCT ATACCATATA ATCAGGAAGC ATCTTAAATG TATAAGAAAA CGCCTCAAAC 720  
 15 CTAATAAAAT GTGTCAATAG CATGTTTAGA ATTAATAATTA AATCTAACA TTCAAGACAT 780  
 TTAATTAAGT AAGGGCGTTC AATATTAAAA TGAACAATGA CTCGTGTTGA AATCATATAT 840  
 20 CATAAAATTA TTTTATAAAC CTTTGAAGAA TACCAGTGT TTTTAGAGTA ATTAATAAGA 900  
 AATAACTTAT AATAGATCCG ATAGCACTTG aGACTATGaa CGTAATCATT AACGGTTTAA 960  
 TGAAGAAGTC TTGAAGCCCA AGGAATATG CTAATGGcAT aCaAATTAAA cTTCcGATGA 1020  
 25 CaCCAGTTCC aAGTACTTca CGACCGCGG CcATAAATAT ATGTTTACGA TATnygTAAA 1080  
 ACATACTAGC CAATAAAACT CCAATCATAC TACCCGAAA TGCAAAAGst GTACCAGTAC 1140  
 CAAAAGAAC TCTTAAATTT GATGATATAA GCGCTTGAGC TAATCCATAC CAAGGACCTA 1200  
 30 CTATGACCGC ACTTAATACA TTTACAAAAT GCTGTACTGG TGCTGCCTTA ACTGGTCCTA 1260  
 GAGGAATGAT GATAATACTG CTTAATACAA CATTATTTGC AATTAAAGT GCAGTTATAG 1320  
 CCAGTTTCTT TGATTTCATA TGATTGTTCT CCTTTTGTG TGTAAATTAAT CACTATGCTT 1380  
 35 GGCCTTATTA TGGTCATTIA AACGTGTTTC CATTGTTGAT ACAAACATTT TCAATAATTG 1440  
 AITCGCTTCA TATTGTGAAG TTTGAAACTG TTCAACTATG GGCAATGTAT TTATTCTGTC 1500  
 40 TTCTATCTC TGAATGGcAT GTTCCGACTG nTCCAGCGCA TTTGTTTCCC GnaA 1554

## (2) INFORMATION FOR SEQ ID NO: 589:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 638 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 589:

TTTGTACGAA AAAAGTTGAT TTACAAATAT ATAAACGTTG TGATTTCAT GTTTGTATAG 60

EP 0 786 519 A2

5  
 10  
 15  
 20  
 25  
 30  
 35  
 40  
 45  
 50  
 55

TATATACATG ACAGCAACTT GGGAAAAAAA GGAAGGTAAC GAAGGTTTAT TAACGTGTAC 180  
 TGTTCTCGCA GAAAAAGTAA ACAAAAGCTC GAGAGATCCC CTCATAATTT CCCCAAAGCG 240  
 TAACCATGTG TGAATAAATT TTGAGCTAGT AGGGTTGCAG CCACGAGTAA GTCTTCCCCTT 300  
 GTTATTGTGT AGCCAGAAAT CCGCAAAACT TCCATGCCCTA AGCGAACTGT TGAGAGTACG 360  
 TTTCGATTTT TGACTGTGTT AGCCTGGAAG TGCTTGTCCC AACCTTGTIT CTGAGCATGA 420  
 ACGsCCGCAA GCCAACATGT TAGTTGAAGC ATCAGGGCGA TTAGCAGCAT GmTATCAAAA 480  
 CGCTCTGAGC TGCTGTGTCG GCTATGGCGT AGGCCTAGTC CGTAGsCAGG ACTTTTCAAG 540  
 TCTCGGAAGG YTTCTTCAAT CTGCATTTCG TTCGAATAGA TATTAACAAG TTGTTGGGT 600  
 GTTcGAATTk CAACaGTaa GTtAgTtGCT AGanCCCA 638

(2) INFORMATION FOR SEQ ID NO: 590:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1242 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 590:

5  
 10  
 15  
 20  
 25  
 30  
 35  
 40  
 45  
 50  
 55

AAAATATTCC CGTACATTTT GATGTCTGTA GGGGCTTTTT TGACTTTAGG ATTTGTCAAT 60  
 TTTTCAATTC ATAAAGGGAG ACGAACGAAA AATGAATCAG CACGTAAAAG TAACATTGTA 120  
 TTTTACTAAT TATAATTACG GCACATATGA CTTAGCAGTA CCAGCATATT TACCGATAAA 180  
 AAACCTAATA GCTTTAGTAT TGGATAGTTT GGACATTTCa ATATTGATG TCAATTACACA 240  
 AATTAAAGTG ATGACGAAAG GTCAATTACT TGTGAAAAT GATCGACTCA TTGATTATCA 300  
 AATCGCTGAT GGAGATAATTT TGAAGTTACT ATAGGAGGGA AAATAGATGG TTA AAAATCA 360  
 TAACCCTAAA AATGAAATGC aAGATATGTT AACGCCTTTA GATGCTGAAG AAGCAGCTAA 420  
 AACAAAATTA CGCTTAGATA TGAGAGAGAT TCCTAAGTCT TCAATTAAAC CAGAACATTT 480  
 TCATTTAATG TACTTATTAG AACAAcATTC TCCATATTTT ATAGATGCTG AATTAACCTGA 540  
 ACTACGTGAC aGTTTCCaAA TACATtATGA CATTaATGAC AATCATACAC CTTTtGATAA 600  
 TATTAAATCA TTTACTAAAA ATGAAAAAAT ACGTTACTTA CTCAATATCA AAAATTTAGA 660  
 AGAaGTAAAT CGTACACGCT ACACATTGTG GTTGGCACCAG ATGAATATAT TTTTCACAAG 720  
 AGATGGATTA CCCATTGCTA AAACAAGAGG GTTACAAAAT GTTGTGTGATC CATTACCTGT 780  
 GTCAGAAGCT GAATTTTTAA CAAGATATAA AGCGCTGGTT ATCTGTGCAT TCAATTAGAA 840



## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 592:

|    |  |      |
|----|--|------|
|    | TTGTTATAGT GTCTGAATCT GTGTAACCTCT TGCCCATGTG TTCTGAGTAA AGCACCACCT | 60   |
| 5  | GTTTATTTAC TTTTCGTTGT AGCTAGCTT CGTGATAGTAG TTTGTTTAAAC TTTTGTGCTG | 120  |
|    | ATATACCGTA GTCTGCCGCG ATTTGAGTTG TGGCTAATGT GCCAGTTGAT TTTAAGATTT  | 180  |
|    | CATCAACATA ATCTGCTTTG GGTTTTAGCT CTCCAATTCT TTGTTGTAAA AGTAAGTTTT  | 240  |
| 10 | GCTCTTTTC TTCTTATAC TCAGTCAACA CTGTAATGAT GTAGTCTGGA TCTTTTAATG    | 300  |
|    | TTTGTTCAAT TACATTGTCT GTTCGTATA TACCGTGTTC GCGAATAGCT GGTaGGACAT   | 360  |
| 15 | CTGATGTTAC CCAGCGTTTG AATTTTCTAG CGGTTTCTCT AATTTTTCG TTTTGTCTTT   | 420  |
|    | GTTTGAAGC ATCAAGATT AGACTGTATA ATCTGATTCT GTTGATAATG ATCATATTTT    | 480  |
|    | TGTTTTGACC TGATGCACTA AATTGGTGGC TCAGTTTGTCT CTCGCTATCA ACATGATTTT | 540  |
| 20 | TAATGGCATT GTCTGATCTT GCATATCCTA AAATCTCAGC AATATCTTTT CCTACAAAT   | 600  |
|    | AAGGTTCTGT TTCAATTCTT ACTGTTCTTA CTGGTAGCTC TTAAAAATTA AATGTTTGTA  | 660  |
|    | ATGCTTGCAT TTGAGTATCC TCCTTTTTC TCAACACCCA CATTGAGTAC ACGGTTATCG   | 720  |
| 25 | CAATGACTAT CGAATGTATT TAAACGCGGC TCATATCATC GCCAGCTCTC GCTCACATCT  | 780  |
|    | GCTCAATGTG GATGTTGATA AGCGTGGTGA TATTAGAAGT TGAATGTTAC TGATTCACCT  | 840  |
|    | TCCGCCACTC TGTTAAATCA GTAACTTTGT TATCGCTTTC AACACCGTTA AGCTTGTCTA  | 900  |
| 30 | ACGCTTTCAC TACTTTTGG AACTCTTGA TAGCACTCG TAGCTTTTGA GTAATTCAT      | 960  |
|    | CTTCTACCAT TTCCAACCA GCAAATGCGT CTTCGTTATT CATGCTTAGA TGTTTGTGA    | 1020 |
|    | AAAGATCTCG AGTGTATCTT ATTCTTTAA GTGATTATC ATAAGCTTCA ATTTGCTCTG    | 1080 |
| 35 | AAAGGTTATG ATATTTTAGT TGTAGTTTAA CTAAATTTAA TGATTGGTCT TGCATTTGTT  | 1140 |
|    | ATGCTCTCTT TAAGATGTTT GTTTCGTTT CGTGACTTTT GTGGGTAAAA AAATATCTCC   | 1200 |
| 40 | AATATTTTTC TCAAAAAAT CAGGATAAT AAACATCTCA TCATTCTTAA ATTGATGCTT    | 1260 |
|    | TCCTAATTCT TTAAACGAT AACCTTCAGT TGATATATTC AAGAGGTTTG CTAATCTTTC   | 1320 |
|    | TTGAGTACAC TTCTTTCTT TTCTCAACTT TATTAAATTC CATTGCATGT TGTCACTCC    | 1380 |
| 45 | CGCTTACAA AACCTACTAT ACAOGATACG GTACTTGTGT CAACATAAAA GTTTGCTTTT   | 1440 |
|    | CGTGATTT   | 1449 |

## (2) INFORMATION FOR SEQ ID NO: 593:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1170 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double



## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 593:

|    |   |      |
|----|---|------|
| 5  | ACCAAAAACA CTTTCTAAAG AAATGTTTC AGAACCAITTT GCTAAAAAAC AATTACGTGA | 60   |
|    | AAAAGCTCGT CTAACCTAACA TTAAGGTTT AGAGATTCCA AGAGTTTCAT ACACTTTAGA | 120  |
|    | CGATGATGAT TTCATTACAG ACGTAGAAAC AGCAAAAGAA TTAAAGCAA AAGGTGATAC  | 180  |
| 10 | AGTCAAGTTC ACTACTAATA AATTCAAAGT ATTTGCTGCA ATTTAGATA CTGTAATTCA  | 240  |
|    | TGGATCAGAT GTAGATTTAG TAAACTGGGT TGA AACGCA CTACAATCAG GATTAGCAGC | 300  |
|    | TAAAGAGCGT AAAGATGCCT TAGCAGTAAG TCCTAAATCT GGATTAGAAC ACATGTCATT | 360  |
| 15 | TTATAATGGA TCTGTTAAAG AAGTTGARGG AGCAGACATG TATGATGCTA TTATTAAAGC | 420  |
|    | TTTAGCAGAT TTACATGAAG ATTATCGTGA TAACGCAACA ATTTATATGC GATATGCAGA | 480  |
| 20 | TTATGTCAAA ATTATTAGTG TTTCTTCAA TGGAAACAACA AATTCTCTTG ACACACCAGC | 540  |
|    | AGAAAAAGTA TTTGGCAAC CAGTAGTATT TACAGATGCA GCAGTTAAAC CTATTGTGGG  | 600  |
|    | AGATTTCAT TATTTTGGAA TTAACATGA TGGAACTATG TATGACACTG ATAAAGATGT   | 660  |
| 25 | TAAAAAGGCG GAATATTGTG TTGTATTAA CAGATGGTAT GATCAGCAAC GTACATTAGA  | 720  |
|    | CAGTGCAATC AGAATTGCAA AAGCAAGA AAATACAGG CCATTACCA GCTAAGCCCC     | 780  |
|    | AAAAGGTTAA TGTAACAGCT AAGGCTAAAT CAGCTGTAAT ATCAGCCGAA TAGGGTGAT  | 840  |
| 30 | GAAATGAGTT TAGAAGAAAC TAAATTGTGG TTGAGAATTG ACTATAATT CGAAATGAT   | 900  |
|    | TTAATTGAAG GTCTCATTCA ATCGGCTAAG TCTGAATTAC TATTAAGTGG GGTnCCAGAT | 960  |
|    | TATGACAAAG ATGACTTGGA ATACCGCTT TTTGTACAG CGATTAAATA TATCATTGCA   | 1020 |
| 35 | AGAGATTATG AAAGTCGTGG ATACTCAAAT GACCAATCTA GAAGCAAGGT GTTTAATGAA | 1080 |
|    | AAAGATTGTC AAAAAATGan TTTGAAATTA AAAAGTGGT AGGTGATTTT TAAATGGAAT  | 1140 |
| 40 | TTAATGAATT TAAAGATCGC GCGTATnTTT                                  | 1170 |

## (2) INFORMATION FOR SEQ ID NO: 594:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1393 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 594:

|   |    |
|---|----|
| TGAAGGCTGT TTTGTTATG GCCAATGCT GTGTATTAA AnCGAATGTT TCGAATGAA | 60 |
|---|----|

ATCATTCCGG AAGAGGACAT GTCCTCAGCT TCTGATTGA AAGCACAGTT TTATCGCGGT 180  
 ACTTTGCAGC GCTTTTACCA ATCnTTGCAC GCAGAAAAGC TTACACCTTA TGTATTGTCT 240  
 5 TATGACGATA TCATTTTCATT TTGTAAGAA AACAAATATCT CTGAAGTAGT GACTGCGGGT 300  
 GATATTATGA GTTATCATCT TGAAGAATAT GATATTTTAC ATCAACGTTT TTTATTCAAT 360  
 GAAGCAGCGA TTGCCGTATC TTTGATACGT GGGAAATCATT ACTTTAAAGC GAGTAAAAACA 420  
 10 ATGAATCAAC AAGGGGAGCC ATACAATGTT TTTACTAGTT TCTATAAAAA ATGGCGACCT 480  
 TACTTGAGCG ATAGAGACGT ATATCACTAT GATTTAAAT CAITCGAAAA CTTTGTCAIT 540  
 GCATCACCTG ATGATTTAGT GTTTGATGAC ATAGCATTTG GATCCTCACA aATAATTGAA 600  
 15 CAGatAAATG GCAACATTTT TTAGATCAAG ATATACAGAA TTACGAAAAG GGAAGAGACT 660  
 ATTTACCTGA AGTATTAACA AGTCAGCTAA GTTGTGCTTT AGCATATGGA TTATTAGATA 720  
 20 TTATTGAAAT TTTTAAATGAT TTATTGGCGC GTTATGATGA AGATGAGGCA AACTATGAAG 780  
 CATTTATAGC TGAACCTCATT TTTAGAGAAT TTTATTATGT GTTAATGACA CAGTATCCTG 840  
 AAACCTCATA CCAAGCTTTC AAACCTAAAT ATCGACAGAT AAAATGGTTCG CAAAATGAAG 900  
 25 CGGATTTTAA TGCATGGTGC GAAGGGCAAA CAGGATTTC CAAATCTGAT GCAGCAATAA 960  
 TGGAATTGAC ACAAACTGGT TTTATGCATA ATCGAATGAG AATGGTTGTG TCGCAATTTT 1020  
 TAACCAAAGA TTTATTATATA GATTGGACAT GGGGAGAAAA ATTCTTTAGA AAGCACCTTA 1080  
 30 TTGACTATGA TGCAGCATCA AATATTCATG GATGGCAATG GTCTGCTTCT ACAGGTACGG 1140  
 ATGCAGTGCC GTATTTTAGA ATGTTTAATC CAATAAGACA GAGTGAACGC TTTGATGCTA 1200  
 AAGCTTTGTA TATCAAAACA TATCTTCCGA TTTTAAATCA AATTGATGCA AAATATTGTC 1260  
 35 ATGATACACA ACGCAATGAG TCCAACCTTT TTGAACAGGG GATTGAATTA GGTAGTCATT 1320  
 ATCCAAGACA AATGGTAGAT CATCAAGAAA AACGTACACA AGTTTATGCT ACATTTAAAG 1380  
 40 CGCTAGACTA ATT 1393

## (2) INFORMATION FOR SEQ ID NO: 595:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1484 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 595:

TTAACTTGAA AGATTCCCC GACATCTATA ACTTGTATAA TGTGTATATG TCGTTTTTAG 60

|    |  |      |
|----|--|------|
|    | AAAAATAAA AAAAGAATAT CTGATTGGT ATAATCAGAC CTIAGAGTTT TCTAATTTAT    | 180  |
|    | CAACAATGT AGTAAGWATA GATACTCCTT TTAAGATAA TCTTTAGAT AATTTAATAA     | 240  |
| 5  | TTTACGCTTT ATACGATCAG TCCAGAGACA TGATTACACT GACAGATGAC GGCCTACTA   | 300  |
|    | TATCTGATTT AGAAAAAAT GGTATTTCTT TAAATAAATC AAAAAACGT AAAAAGATTT    | 360  |
|    | TTGAAGAGCA CCTTCAGCT TACGGTATTA AATATAACGA TAAACTCAC GAAATTTTGT    | 420  |
| 10 | TTCAAATAA CTTTAAAAAT TTTAATAAAT CGAAACATAA TTTATTACAG TGCCTTATAT   | 480  |
|    | TTGTTAATGA TATGACTTA CTTCTAATC CTAAGTCACA GAACATATTT ACAGAAGATG    | 540  |
|    | TTGCAACAA ATTGGATGAA CATACATTT ATTACGGAAG AGATTTACCT ATTATAGGAA    | 600  |
| 15 | GCAGTGGTGT TGTTCAATAT TTCGACTTTT TTATTAGCGC TAAGAAAAAT CAAAAAGAAA  | 660  |
|    | AATTTATCAA TGCTATTTCT AACCTAATA ATTCTATGAT TATTAAGTCG AAAATAACGG   | 720  |
| 20 | ATGCTATGCA AGCAAAAAAA ATAAAAAGAC ACAGGCCAAA TGAGTTTATT TTTATTTTAA  | 780  |
|    | ATGACTCAAA AAAAGAAATA AATGAACATA ATAAAAATCT TCTTCATGAA AACTATATTA  | 840  |
|    | GTACAATAGA TTATAGCGAA TTAGATGAAA AGATAGGTTT ATTGATTTAA TATATATAGA  | 900  |
| 25 | CGTGATATA TCAATGTTTA TATTAAATGA AACGAATTTG AAAATTCGA AACAGCTTAA    | 960  |
|    | GATAGCAAAAT TGAATAGCCT TATTGATAAT GCAGAATCGT CTACACTTAG TTGAACAAAT | 1020 |
|    | TCTATGAGAA TAGATAATTGT TAAACTATTT GGGTAGGCGA TTTTATTTTG ACAAGAGTCA | 1080 |
| 30 | GAAGATTATT TAGTTCATAG TTTAAGTTAC AAATGGCTAA ACTTTATGAA AATGGTATGA  | 1140 |
|    | CTAGAATCGA AATTATACTC GAATACGATT TAACACTCTC AATCTTCTCG AACTGAATAA  | 1200 |
|    | ATACTGGGTC ATTCAATCAT CAAGACAAC TAAATAAGTGA TGATAAATGT TCATACCATT  | 1260 |
| 35 | TTGCAAAATAT AACCAGATAC TCTCAACAAA ATGTAGAGTT CAAAATTTAA AATGTCCTTA | 1320 |
|    | ATTGGCACAT TGCTGTATGA TAATAAATTC AACTTAGATT TCAACAATAT AACTAATGAT  | 1380 |
| 40 | GCTTATATTT ACGCGATAT CAAACCAAAC AGCCCATCCA CAAACAAAT AGAATATATC    | 1440 |
|    | ATTGTTAATAA TAGTAATGA TATATCTTCC AAAGACCGCA AACT                   | 1484 |

(2) INFORMATION FOR SEQ ID NO: 596:

- 45 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1435 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 596:

55

# EP 0 786 519 A2

CTTATCaTT AAATGCCaTT TTGaATCCITT CTtGAATTtG TTTGATTtCC TTTTCAGTAT 120  
TATTAAcCCTT GTtCAAGACA GTGCCATTAA TTTTCTtCtT GAGtGGACCA ACCTCaTCTT 180  
5 TATTAGtTGA TTCAtGAATC GTtAAATCCAA TGtGTGGCTG TATTCTATAA CCATCaTTcG 240  
CTATAGtTGA AACATATtGT GaTAATTGTa ATGGtGTATA GGTATCaTAT TGACCAATtG 300  
10 ATAAATCTAG ATAATTACCT GGATtATTtG TTAATGGtTC GATTtGACCT CTtGTtTcAT 360  
TTGGTAAATC TATCCCTGtT TTCAcACCTa AGCCtACTtG ATtTAATCCCT CTtCTTAGCT 420  
TTTGGGCAGG TGACTtTATG TCTGAAGGTa AAGCCATACC AGAATAATAA GGGTCTCCCG 480  
15 CTAAATTTAA TGCTGtTTTA AACATATATA CGtTTGATGa ATGCATCAAA GCTTGTtTAT 540  
CATTAAATAGT TACATGCCCG TTTTtATTGA AGTATGATCG TTTTGTCAAA CCACCTtGGA 600  
AATGTAAATGG TTCAATGACC ATTGTtTCTC CAACtTTGAT AGCTtTATTC TGATAACCGG 660  
20 CTAAATATGT TCCACCTTTT ACAGAAGATC CAACCGCAAA TTGAGAAGTa AACGTACCAa 720  
TGtCATAAATC AGtCaTTtTA CCACTCTTAT TAATCTGCTT TCCGGCAAGC GCAAGAATGT 780  
CTCCATTTTt AGGATTTtGT ACAACCATCA tTGCATTATC CATATCTtTG GCACCTtGAC 840  
25 TGCGaSTtCT TAATTtGtTT ATCTAAATaT GCTtCTACTT CTtTTtGAAG ATCTATATCG 900  
ATOGtTAATt TCAAACTCTG ACCGCGAGCG CCAAGtTTTA ACaCTTCTGa AGATGTAACT 960  
TTACCAGaTT TGtCCGtTGT GTATTtCaTT TCTtTCTTCT TACCACGCaA TACATCTTCA 1020  
30 TATTGATATt CTAGGTAAAG TTTTCCAACA CGATCaTTGC GTGAATATCC TTTGGATAAG 1080  
TAATGTtCTG TCAATtCTtT TGGaATACCT TCAGCAGGTG TCGATACATC TCCGAATATA 1140  
CCTCTtAAAG TATCGCCATA TGGATATTTT CTATCCCAAT CCATAGACGT GTTAACACCT 1200  
35 GGTAAATtTG AAAGtTGCTG AGAAACTGCT GCATACTCTT TtTCACTGAC ATCTtCATtT 1260  
TTTATCaTTt GTGATCTAA AACTGTtCCT GGCAtTCATC TCTCGAAAAA TAGCTAAAAC 1320  
40 TTGGTAAATC TTTAGAAGAC AATTCATCTa ATTGTGGATT TTCCGGATtT CGGATAACAG 1380  
TTGGtTtATC CATAATGGAT CCTGGtTTAA TACCTtCCaA CnGGCGAAAC TAGCC 1435

(2) INFORMATION FOR SEQ ID NO: 597:

45 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 516 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear  
50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 597:

55

AGAAAAAGAT TTAGAAAAGA TTACCATTCG CGACATAACA ACACGCGCTG ATATCAATAG 120  
 AGGGACATTT TACTTACATT ACGAGGATAA ATATATGTTA CTGCGAGATA TGAAGATGA 180  
 5 GTATATTTC AACTAACGA CATATACTCA ATTTGATTGG TTACGCGGTT CTTCATTGTA 240  
 AGACATTGCG AATACTTTTG TAAATAATAT ACTCAAAAAT ATTTTCCAAC ATATTCATGA 300  
 10 TAATTAGAG TTTTATCATA CTATTTTACA ATTGGAACGC ACGAGTCAAT TAGAATTGAA 360  
 AATCAACGAA CATATTAAAA ATAATATGCA ACGTTACATT AGTATCAATC ATTCTATCGG 420  
 AGGCGTTCCA GAGATGTATT TTACAGTTTA CGTTCCGGA GCAACAATT CaATTAnnTA 480  
 15 AATACTGGGG TAATGGACAA ACAGCCCAT TTCAGT 516

(2) INFORMATION FOR SEQ ID NO: 598:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1955 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 598:

AGTACTTTAC CATGAGGTAA GACTTTTATC CCTGCTTTA TAGATTGTGC TTTTATTTGT 60  
 GTGCCAGAAC CAATAGTact GGTCTAATGA TTTTGCACTC ACATTATCAA TAGCTTTTGT 120  
 30 TTTGCTTGGA CCTTCAACAT CTGGCACTTT ACTATACTCT GCATTAGATG TGTCACTCTT 180  
 TGATTTACCT ACATTTAAAT ATTTCAAAGT ATTTTCCATT ATTGGTTTAA ACGCTTTACT 240  
 AACACCTAAT TCATAAGCTT CTGGGTCAAT TTTTGTGCCC AAGCTCATAC CAGCGTATAC 300  
 35 AATAACTTTA GGATTTTICT TCGGCGCGTC ACCCATAAAA CTTACAAAAT ATGGGTTTGG 360  
 ACCTTTAAGC TATCAACAC CATTAGGTGC AGCGACTTGT GCTGTACCAG TCTTACCTTC 420  
 40 GACCTCATAA CCATCAATAC GATAGTTTGC AGCGTGACTC TTCTTACTAT TCACAACATA 480  
 ATCCAATTGC TTTTCAACTT TTTCAAGCAGT ATCTTTTGTT ATTGGTTTGC CTGCGATTGG 540  
 TTTTGGCCTT TTATAAAATT GTCTTTTACT AACAGGATT TCAACGCTAT TCACAACCA 600  
 45 TGGTTTAAAC ATATTACCAT CATTAAAGAA CGCTGATTGC GCTTGTAAAC TTTGAACAGG 660  
 TGTACTGTT GTCGATTGAC CAAATGATGA CGTTTTTTG TGCAACTCAT TACTCCATCC 720  
 AATTGACCA GGTGCTTCTC CATCAACAAT ACCTTTAGTT GaTTTTCCaA ATCCaAATCG 780  
 50 TCAATACCA GATTTCATTT TGCTGCAACC AACTAAATCT TGTAAATGCA TCATCAATGT 840  
 ATTAGATGAA TAAGTAAATC CGAGTGACAT TGGGATTTCa CCCCACCGA CTCATTCCa 900



EP 0 786 519 A2

TATTAGTATT AGGGTTATGG TATTTTATAT TGCTAGCAAG ACTTCCACCA ATTAATCCAA 480  
 GCCCAACAAA TAAACTGTGT GTCATATAAA TCACCTTATT TCGAAATTTT CAGAATAATA 540  
 ACATTGTAA TGAGCTGTGT ACACAGTGCA ATAGTAAATA AAAATCGATA ATAGCATTAA 600  
 TAGATAAACG GAGATAAATC ATCTACAATA AAGAGTATAG TAACACAATG GCAACGGAGG 660  
 GGTAAATCAA TGGAACCAAT ATTAGAAATG ATTAAACAT TAAC 704

(2) INFORMATION FOR SEQ ID NO: 600:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1158 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 600:

CCTCTAAAGG ATCAGCAACC TTTTTCATCG TGACTAAGCC TATAATACCC TTAACCTAT 60  
 TATTATTAAC TTTTACCTCT GTGTATTCCT TATCAATTAG GCGACGCCAG TGACGTTTAT 120  
 CTATATACTT TACTTTCACA GTACCAACT CCTTGTCATT ATTATATAAA ACTACATAAA 180  
 TGATGTCATG TCATGATACA GTTTTATAGT TTTTGTAAAT CATATGAAAA TATTTATTAT 240  
 TATTTTACTA CAACTCGCTT CAATTACTT AAAATAGACA ATATTAATTA GATAGTACAC 300  
 ACATTTCTTC ATAAAGTGA TTTTCAAAA ATATAAATA CACACTCTTA TCGTTTCAA 360  
 AATCAITTTAA TGCTATTTC ATTAATAACA GCTGAAGCAT CAAATCTATT CTGATTCAAT 420  
 CAAGAATACA TATAAAATG AAGTGACTCA AAGGTTTATT AACCACTTC AAACCACTTC 480  
 ATTGATCACT TTTATTTTAA AGCATATTT TCGATTACTG ATTTAAGATG CGGATATTGC 540  
 GTAATTAAAT CAGATTGCTT AAACAATTCA AACTGCTTAA ACTCAAATGC CGGTGACAC 600  
 ATACAACCTA CTAAACTAAA AGTATTTGAT ATTTGCGATT AAGAAGCAAA AATTGTTCTT 660  
 TTAGGCACTA CATATTGCAA TACATCTCCA TTTTGGATAT CAGTACCCAA TGTGTCAGTC 720  
 GTATATTCCC CATCCGATT TATCATATGA ATTGTTAGAG AATCGCCAGC ATGATAGTAC 780  
 CATACTTCAT CAGCATCAAT TCGATGAAAA TCGCAAAATAT TGTCATCTGT AAGTAAAAAA 840  
 TAAATACTAC TAAACGGCGC TCTGCGTCCA TCCTTCAATA CTCTCGAAT TGTCCTCTTA 900  
 TAGAAACCAC CTTGAGGATG TGATTCAAGT TGCAATTCAAT CACTCCGATT TTCTGCTGAT 960  
 TTCATTATTT CAAATCCACA TTATGGAATA GTTTTGTAC ATCTTCTAAT TCTTCTAATG 1020  
 CATCGATTAA TTTTTCAAAT GTTACTTGGT CCGCTTCAGA AAGTTCAATA TCTGTTGAG 1080

GAACGACTGC AAATTGAT

1158

## (2) INFORMATION FOR SEQ ID NO: 601:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 827 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 601:

|    |   |     |
|----|---|-----|
| 15 | TGATACTTAT TTTTCAATAT TATTTGGAAT AAATCTTnT AATTGTGGTG TGAATAAAAC  | 60  |
|    | ACCGAAATTC TTTTCAGATT CATCTTTTcC ACTACCTGCA TATAAAATAC CAATCAGTYC | 120 |
|    | ATGTTTAGAA TTCAATACAG GAGATCCTGA ATTACCTGGT TGTGCATACG CATCAAATTC | 180 |
| 20 | CATAAACGTT CCACTGATAT GGTTAATCGT TCCTGTCGAT TCAACATTT TATATTTTGT  | 240 |
|    | TGTGTCACCC TTGCGATAAC CAATAACAGA AATTCTATCT TTCACCTTTG CTCGCTCTGC | 300 |
|    | AAATTTTGTA TAACTAACGT TCTTATTAAA ATTCAAACCT TCTGTACTTG TTTCATGAAC | 360 |
| 25 | ATGAACTATC GCAAGGTCCT CTTTCCGGG ATATTCTACA ATGCTTTTAA COTCGTAGTT  | 420 |
|    | TCCTCCGCCT TTACCTTTAC TCGAATGATG TGCATGACT CTATTTTAA AATATCAATT   | 480 |
|    | ACTTTTAGCG ATATGTTTGT TAGTTACGAT TGTATTTTAA CCAACAATA CACCAGTACC  | 540 |
| 30 | ACCCACAAAT GCTACCACIG aATTGTATGG TTCTTAGTT GCATCGGTAA TTTCTTTGAC  | 600 |
|    | ATTCTTTTCT GCTTTGGCAA TTGAATGArG CTGATTAGAA ATATTTTCAG CAAAACCAAG | 660 |
|    | AGATGTTAAA ATAGTTAAAG CAGTTAAACC TTTAACCAAT ACATTTTAT TCATATTTTA  | 720 |
| 35 | CCTCCTTCCA AATTTATTTT nATAGTTAAT GTATCTAGAT ACACATTTAC TATATACATT | 780 |
|    | TTAAATAAAA ATATTnATTA AATATTAATT AACCATTAT nAACGTT                | 827 |

## (2) INFORMATION FOR SEQ ID NO: 602:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 796 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 602:

|    |  |     |
|----|--|-----|
| 50 | nGnACGTAA AGTATTTATG TTAATGGCA nAAATATATG AATGATGAGT TACCTGGATT  | 60  |
|    | ACCAATGTTc CAAGGTAAT CGATAACTAT TGTTAACGAT AAAGTACGAA ACTTAGACAT | 120 |



EP 0 786 519 A2

|    |            |            |              |            |            |             |     |
|----|------------|------------|--------------|------------|------------|-------------|-----|
| 5  | CGACGATTTT | GTCGAGGTTT | ACTTATATGG   | GGGGATGTTA | TGAATAATGT | ATTGTTAGAG  | 240 |
|    | GTTAAGAGTT | TAGAAACATC | ATTA AAAAATA | AATAATGAAT | GGTTAGCAAC | TGTTGAAAAT  | 300 |
|    | ATTTCCTTTG | AATTATCTAA | AGGAGAAAGT   | TTGGGTATAG | TAGGGGAAAT | TGGTTCGGGT  | 360 |
| 10 | AAGTCCATAT | TAAGTAAGTC | AATTATTAA    | TTATTACCAG | AAAAGATATC | TAAACTAAGT  | 420 |
|    | AATGGAGAAG | TTATATTGTA | TGTTAAACGA   | ATCGATACGC | TCAATGAGAA | GCAATTGCTA  | 480 |
|    | GATATTCGAG | GAAATGATAT | TGCTATGATT   | TTTCAAGAAC | CTATGACTGC | TTTAAATCCT  | 540 |
| 15 | GTATTTACCA | TAAAAAATCA | ACTTGTGGAA   | TCTATAAAAT | CACATAAAAA | AATTTCTAAA  | 600 |
|    | AAAGAAGCAA | ATAAAATTAG | AAAAGATTTA   | CTAAAAAAGG | TTGGAATTGC | TAGACAAGAT  | 660 |
|    | GAAATATTAA | ATATCATATC | TCATCAATTA   | TCTGGTGGTA | TGAGACAAAG | AGTAATGATT  | 720 |
| 20 | GCAATGGCCA | TTTCATGATG | TCCTAAATTA   | TTAATTGCTG | ATGAACCTAC | AACAGCATTTG | 780 |
|    | GATGTCAAGA | TTCAAG     |              |            |            |             | 796 |

(2) INFORMATION FOR SEQ ID NO: 603:

25 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1618 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEO ID NO: 603:

|    |             |            |            |             |            |            |     |
|----|-------------|------------|------------|-------------|------------|------------|-----|
| 35 | GGATGTGTTTC | ATCCCAAATT | GaAaCgAaAa | TTTATGATCA  | AATGATGTGG | ATTGGTITTA | 60  |
|    | AaCCATTCAa  | AATTaCCAAT | ATTAmAaAG  | AATCTGAAGA  | CATTAAATCA | TTTtACAGTT | 120 |
|    | GAACtGAAG   | AATATGACTT | TAGTGAATTT | ACACCAGGCC  | AATACATCAC | AGTtGATGTT | 180 |
| 40 | TCTAGTGATA  | AACTTCcATA | TAGAGCTAAa | CGTCACTATT  | CTATCGTATC | AGGTGAAAAA | 240 |
|    | AACCATTTAA  | CTTTTGGCGT | TAAACGTGAT | GTCACAAcAG  | AACATGAAGG | CGAAGTTTCA | 300 |
|    | ACAATTTTAC  | ACGATGAGAT | TAAAGAAGGC | GATATGATTA  | ATTtAsTGCG | CCTGTAGGTG | 360 |
| 45 | GATTcGTATT  | AGAGAATACG | ACTGAACCAC | AACtTTTCTT  | AGGTTCAAGT | ATTGGTGTTA | 420 |
|    | CACCTTTAGT  | AGCTATGTAT | GAAGCTGCCT | CTGcCAAAAG  | TTTAGATACA | CAGATGGTTC | 480 |
|    | AAGTTGTCTA  | AAATGAACAA | CATTtACCTT | TCAAAGACAA  | CTTCAACAGT | ATCGCAAGCC | 540 |
| 50 | ATCATGACAA  | CGCTAAATTA | TATACACACT | TnAGATATAA  | CAAGCTATAT | TTGGTGCTGA | 600 |
|    | AGAATTACAA  | GTATTtTTAG | CAATAAACC  | TGAaATTTAT  | ATCTGTGGTG | TGACAAAAAT | 660 |
| 55 | CTTACaaTCT  | ATGATCGAAG | CACTTAAATC | TTTAAATTTAC | GATATGGATC | GCGTACACTA | 720 |

|    |            |             |            |            |            |            |      |
|----|------------|-------------|------------|------------|------------|------------|------|
| 5  | ATATGAAACT | TAATCAACGT  | TACGTAAAAG | TATTTCGATT | ATATTTCGTA | AGTATTGTTA | 840  |
|    | CTGCAATAT  | TATTGTTAAA  | AATAATAATT | TAATTAAAC  | ACTGATACAA | ACCATAGCCG | 900  |
|    | GGTACACGGT | CTTTGCAGTT  | GGTTTGAAGT | ATTTAACATA | ACGTAAAAAT | AAATGACATC | 960  |
|    | TATACCAAAA | CAGCTATGAC  | TTTACTTTGT | CATGGCTGT  | TTTTAATTGG | GAGTAGGACA | 1020 |
| 10 | GAGATGATAT | TTTCGCAAAA  | TTTATTGTGT | CGTCCACCA  | CAACATGCAT | TGATGTATGC | 1080 |
|    | TCACTGAATT | TCATAAGAAA  | GGAGTTCACA | AGATGACCGT | AGATATTGGA | CGGATTTATG | 1140 |
|    | ACAATAAAGA | TAATACCGAC  | GCTATTTCGA | TCCTAGTCGA | TAGAGTCTGG | CCGAGAGGTA | 1200 |
|    | TTTCGAAAAG | AACGTCTAAC  | CTAGATTATT | GGTTAAAAGA | CATTGCCCGT | TCTACTGAGT | 1260 |
| 15 | TGCGACAATG | GTTCCAAACAT | GATCCTAAAC | TTTTTGGAGC | TTTTAAAGAA | AAATATGAAA | 1320 |
|    | AAGAATTACG | TGATCAGGAT  | GCGCAAAAAG | ATGCTTTTGA | AAAATTAAAG | GATATTGTAA | 1380 |
|    | ATCAGCATAA | TCATGTGCTA  | TGTGTATATG | CAGCAAAAAG | TACTAAACAT | AACCAAGCTG | 1440 |
|    | TAGTACTACA | GCAGTTTCTC  | AATACTTAGT | TATATATTGT | TATCGTGAAT | CACATAAAT  | 1500 |
| 20 | TCGCTAACT  | TCATTTCTAA  | TATTATTATT | TAAAAAGCAT | AATCTACTAA | ATTAATTGGA | 1560 |
|    | ACTTTAATAA | TTAATAAACT  | TGGAACAATA | TTTTACTTTA | ACAAAATAAA | GTTTTAAA   | 1618 |

30 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 2115 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

|    |            |            |            |            |            |            |     |
|----|------------|------------|------------|------------|------------|------------|-----|
| 40 | ATTATAATGC | AAATATTGAA | GCATTTAAAA | CAGTCGCTAA | AGCAGTAGGC | AAAGAGAAAG | 60  |
|    | AAGGCGAGAA | CGCTCTGGAA | AAGCATGATA | AAATATTAGC | GGAGATTAGA | AAGAAAATTG | 120 |
|    | AACAGAGTAC | GTTAAAATCT | GCATTTGCAT | TCGGTATCTC | AAGAGCAGGT | ATGTTTATTA | 180 |
| 45 | ATAATGAAGA | TACATTTATG | GGACAATTCT | TAAATAAAAA | GGGTATTCAA | CCTGAAGtca | 240 |
|    | mAAAArAmAA | AACTACGCAT | GTTGGTGAAC | GCAAGGGTGG | TCCTTATATA | TATTTAAATA | 300 |
|    | ATGAAGAACT | TGCCAATATC | AATCCAAAAG | TTATGATTTT | AGCCACTGAC | GGAAAAACGG | 360 |
| 50 | ACAAAAATAG | AACGAAATTC | ATTGATCCTG | CAGTTTGGAA | ATCATTTAAA | GCTGTGAAAG | 420 |
|    | ATAACAAAGT | TTATGACGTT | GACCGAAATA | AGTGGTTGAA | ATCAAGGGGG | ATTATCGCAA | 480 |
| 55 | GTGAAAGTAT | GGCAGAAGAT | TTAGAAAAAA | TTGCAGAAAA | AGCAAAATAA | AAATACAGCG | 540 |



(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 605:

(2) INFORMATION FOR SEQ ID NO: 606:

(A) LENGTH: 1097 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 606:

|             |            |            |            |             |            |     |
|-------------|------------|------------|------------|-------------|------------|-----|
| GCTnTTATAT  | TTAAAAAATT | TATTnGCGGA | TAATTACCTT | TATTAAATCC  | CACCCATTAG | 60  |
| GAwTAcGAGT  | AATAGGACTA | ATACTGGAAT | CACATATGTG | AACATAACGT  | CCCTCCTTTA | 120 |
| ACTTAATTTT  | AATTGTAATC | AAATTTGACA | ATAAGTCAAA | mCATTAAATAC | CTATGATArG | 180 |
| TATCAITTTAT | TAACATATGT | ATCATATTTT | TAATCTIGCG | TAATTTTITAT | CGTAACTAT  | 240 |

## EP 0 786 519 A2

CTTTTGTCAT TTTAATTATG TTAAGATAAT AGTAGTTATA GAAGTTCAAT CTATAGGAGG 360  
 CATAGCATGG ATATTCCAAA AATCACGACA TTTTAAATGT TTAATAACCA AGCTGAAGAA 420  
 5 GCTGTTAAAC TATACACAAG CTIATTTGAA GATAGTGAGA TTATAACAAT GGCTAAGTAT 480  
 GGTGAAAATG GACCTGGTGA TCCCGGGACT GTACAACACT CAATATTTAC ATTAATGGA 540  
 10 CAAGTATTCA TGGCGATTGA TGCTAATAGT GGCACAGAAT TACCAATGAA TCCTGCGATT 600  
 TCATTATTTG TTACAGTAAA AGATACTATT GAAATGGAAC GACTATTTAA TGGATTAAAA 660  
 GATGAAGGTG CCATTTTAAAT GCCAAAAACG AATATGCCAC CATACAGAGA GTTTGCTTGG 720  
 15 GTTCAAGATA AGTTTGGAGT AAGTTTTCAT TTAGCATTAC CTGAGTAAAA GGATTGCGAC 780  
 AGCTTGAAT GATAAAGAAA CACTTTTCTT TATGCATGCG TTTTACCTAT GTATTTCTAT 840  
 TTTAAGTACA CATTAGCAIT TTAGTTTCGC TCATTTTAA AATCACAAG ATAATTGTGA 900  
 20 ACTTAAATGA TGATACGTTA TTTAAAAACA CGTACTTCG TTTCAATGAA CGCATTAAAT 960  
 AATAAATAAA CACCTCACCA TAAGAAAGGA CTACTTCTT TTGTGAGGTG TTTATTTGTC 1020  
 25 GACAACITGA TTATGATTGT TTCATTTTTT GAATAAGTTC ATAATCAGGT GTAGCATACA 1080  
 AAGTTTTTGG ATTGTCA 1097

## (2) INFORMATION FOR SEQ ID NO: 607:

30 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 2031 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 35 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 607:

40 GAGGTCTAGC ATGTCTCGTT CAAAAAATA CTTTACTTA TCTAGCTTAA TGATTATTTT 60  
 AAGCTTTTTC TTTAATACAA ATAACGTTTT CCTAAGTGGA CTTTTTAATT CTTTATTTAA 120  
 ATTAATACTT TTCTGCAGTG TTATTAACTC AATTGTACTA ATTTTGCTA TAATTTTTCG 180  
 45 AGATCGTTCA ATTAATACAC TAAAGCCTGA TGCAGATTGG ATTAGAATTG CGAGTAAAG 240  
 TTTGCTTTGG aTTATTCTAA TTGTTATTTT AGTACATATC TTTTCAATTG TTCGTACATT 300  
 50 CGGTTTTTAT TAAAAAGTT AATATGTCAT TGTAGCCTTA ATACAAAACA ATACAAATGA 360  
 TCATGCTATA ATGAGTAAAA CAATTGATA ACGTTGTTGC GTATAAAAAA TATTAGATT 420  
 TCGAAATCAT AACTATGCAT CTAATCGCTA TAGTTATACA ACAAGATATA ACATATAATG 480  
 55 AGGTTTGATA ATGCATCGAC AATTTTGTGC GTGCGTTTGC CAAAACCTCT TTTTAAATT 540

EP 0 786 519 A2

|    |   |      |
|----|---|------|
|    | TTTAGGCGCT TTATTACGTT ACCTCATTTT TTTTCTGAAT ACTGACGGAG GTTTTCCAAT   | 660  |
|    | CGGAACACTG ATAGCCAATT TGACTGGTGC CTTTGTAAAT GGATTGCTAA CAGCCTTAAC   | 720  |
| 5  | AATTGCATT TTTTCAAACC ATCCGACCCT AAAAAAGCT ATTACGACTG GTTTTCTTGG     | 780  |
|    | TGCTTTAAGC ACTTTTCAA CATTTCGAAT AGAATTAATA CATATGTTG ATCATCAACA     | 840  |
| 10 | ATTTATAACT TTACTACTAT ATGCTGTAAC AAGTTATGTC TTTGGTATT TGTATGTTA     | 900  |
|    | CGTCGGTATA AAAC TAGGTG GTGGTTTATC ATGATATCAA TCATTTTAGT CATGATTGGC  | 960  |
|    | GGCGGTTTTG GCGCAATTGC TAGAAGTGCC ATTACTGATT ATTTTAATCA TAAATTACT    | 1020 |
| 15 | TCAAAGTTAC CTATCGCAAC ATTGATAGTA AATCTAGTTG GTAGTTTTTT AATTGGATTA   | 1080 |
|    | ACTATAGGCT TATCAATTTC AATCTCATGG TTCCCTGCGT TCTTTGTTAC CGGTTTTTTA   | 1140 |
|    | GGTGGCTTAA CAATTTTCTC AAGCTTAGCC AAAGAACTTA CACTAATGAT GACGCCAAA    | 1200 |
| 20 | TTTAATAITA ACCTTTTCTT CAATTAATCA CTTTACAAT TCATCAATGG ATTTATAGCT    | 1260 |
|    | TGTTATATTG GCTATCATAT TTAATAATAA AATGCTTCAT TCAGCAATAA GGTAAATTAC   | 1320 |
| 25 | GACACCTTCC TGAACGAAGC ATTTTTTAAT TTTTCATGCAA ATTTTTTAAGC ACCATATAAT | 1380 |
|    | GCCTACCAAA TTTCAATAAT CTTTGTGCGC GTTTAAATAA TGTGAATGTC AATAAATTCT   | 1440 |
|    | CCAAACTAGT CGAAAATAAA GGGAGTGGGA CATAAATCCC TAAAAAACA GCAGTAAGAT    | 1500 |
| 30 | AATTTTCAAT TAGAAAATAT CTTACTGCTG TTCTCTATTT ATACAATACT TCGTATTGAA   | 1560 |
|    | TGGCTTCGCT TTCTAGGGT GCCGTCTCAG CCTCGGCTTC GACTGGCACT GCTCCCTCAG    | 1620 |
|    | GAGTCTCGCC ATTAATACTA CGTATTAACA TGTAAATTTA CTTTACATA CTTTAAAAAA    | 1680 |
| 35 | TAGACACTT TGCCCAACTT AACTACCAA TAGAAACCTC TGTTAGAATT CCTCAAAATG     | 1740 |
|    | ATATTTGCGC ATATGTTAAT GAAATTGTTG AAACGATACC TGATAGCGAA TTCATGAAT    | 1800 |
| 40 | TCAGACATCA TCGTGGCGCA ACATCTATC ATCCAAAAAT GATGTTAAAA ATCATCTTAT    | 1860 |
|    | ATGCATATAC TCAATCTGTA TTTTCTGGTC GAAGAATAGA GAAATTACTT CATGACAGTA   | 1920 |
|    | TTCGAATGAT GTGGTTAGCT CAAAATCAAA CACCTTCTTA TAAACTATT AATCGTTTTA    | 1980 |
| 45 | GAGTGAATCC TAATACTGAT GCGTTAATTG AATCTTTATT TATTCAGTTC C            | 2031 |

(2) INFORMATION FOR SEQ ID NO: 608:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 687 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

|    |             |             |            |            |            |            |     |
|----|-------------|-------------|------------|------------|------------|------------|-----|
| 5  | TCCCCGGGAA  | TCGAACCGCG  | GAnCTCTGCG | GTGACATGCA | GGGCTGTAA  | CCGCTACACT | 60  |
|    | ACGAGACCTA  | TTAnATTAA   | AACTATGTAT | TGCGGGAGGC | GGATTGAAC  | CACCGACCTT | 120 |
|    | CGGGTTATGA  | GCCCGACGAG  | CTACCGAACT | GCTCCATCCC | GCCTTAATAT | TATTTTGAAT | 180 |
| 10 | TACCTAAATTA | ATATACCATA  | ATCAAAAACC | TAAAGTrrG  | AACTTTTGA  | ATTTAATTTA | 240 |
|    | AATGTTATCT  | CTTAAATAAT  | TACTTAAATA | TCGTAGCAAC | ATGTTCTCTG | TTGAACACAA | 300 |
|    | ATATTAGTAT  | ATTCAATTTT  | GTAGTGTACG | TCAACGACAT | TTTCAAAGTT | TTTTGTGTAA | 360 |
| 15 | AAAAACGCTT  | CTTATCCCTT  | TTTATCATAT | AAGTGTCTAA | TAGTTGTCAT | AAATAGTGkT | 420 |
|    | AAAGCATTTA  | AAAAGGTATA  | GGAGTTATAA | AGTTTACAAC | GCCTATACCT | TCTGAAAAAG | 480 |
|    | AAATTATAAA  | AGCTTGTGTAC | ACCGCATATT | CTTTCACTCA | GCGACTACCA | ATATAACATT | 540 |
| 20 | GTAGCCCLAA  | GACATTGCTT  | GACCCCTCAN | TTTACAACAT | TTTTCAAAAT | CAGCAGCTAC | 600 |
|    | CTACTGCAC   | AACATACAA   | AACCCnTACA | CTAnCTATCG | TGTCATGTAA | TCTTGCATCC | 660 |
|    | GATCTTGCAA  | CGCTGTAAAT  | TTTTTCA    |            |            |            | 687 |

(2) INFORMATION FOR SEQ ID NO: 609:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 843 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

|    |            |             |            |             |             |            |     |
|----|------------|-------------|------------|-------------|-------------|------------|-----|
| 35 | GTGTTGAATA | CTTTGACAAA  | ATGAAATCAT | TTGTTCTTTC  | GTTTCAAATT  | TAAGTGTTTG | 60  |
|    | AATTAAATCA | GTTCTGTTTT  | CGTAGACTTT | CGCGGTGTG   | TTCATATTCA  | TTTTTCTTAA | 120 |
| 40 | AAATAAACTA | GTAACAACATG | CACCTTTAAG | ACTCTGACTG  | ACAAGGTGTG  | GTGCTAAAAA | 180 |
|    | GAAACCTTGA | TACATTTCAA  | GCAATGCATT | TAATGATGCA  | CCCCTTCTT   | TACCAATACC | 240 |
|    | AGGTGCTGTC | AATCTATAAC  | CACATCGTTC | AATTAATCT   | TTTCTACCGA  | CAATGTATCC | 300 |
| 45 | ACCAATCTTA | GCTAAACAC   | CGCCAGGGTT | TTTAATTAA   | GATCCTGCTA  | TTAAATCGGC | 360 |
|    | ACCACATTTA | ATAGGTTTCA  | GTCCTTCAAC | AGTTTCCCA   | TAAACAGTTAT | CCACAAATAT | 420 |
|    | TAAATATTCT | GATGACAGT   | TTTCTCAACT | AATTAATTAAC | TTTTCATTTT  | CATCTAGCGG | 480 |
| 50 | TAATTGAAGT | CTTTGATCAT  | AGGCTTTTGA | ACGTTGAATC  | TGCTATTACTT | TGGTGGGCTC | 540 |
|    | AGAAACCCCA | TCTAACACAC  | TTTCAATATC | GATCTTACCT  | TCTTTAAGTG  | CAATATCTTT | 600 |
| 55 | ATACGATmCG | CCATGCTCCA  | TTAAACTTTT | AATACCATT   | CCGTTTACGC  | CAATGACTTC | 660 |

## EP 0 786 519 A2

|            |            |            |             |            |             |     |
|------------|------------|------------|-------------|------------|-------------|-----|
| ACTTTGTAA  | GCAATAGTAA | TCGCATCGGT | ACCTGAAATA  | ATTTCGGGAC | GAACAAATTGC | 780 |
| ATCTTCTGCT | TTAAATGCCT | GCGCATATAT | TTCTTAAATTG | AACGAATTTA | GAACAAATCG  | 840 |
| TAA        |            |            |             |            |             | 843 |

(2) INFORMATION FOR SEQ ID NO: 610:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 775 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 610:

|            |             |             |            |            |             |     |
|------------|-------------|-------------|------------|------------|-------------|-----|
| GATTATTATA | AACCGAATGA  | AACAAAAAG   | TTTTACATA  | TTTCAACAA  | CTTTAITTCA  | 60  |
| ATAAAGCATC | AAGTTGATCT  | AAAGTTTGAT  | TCATACCTTG | TTCAACACCC | ATGTTTATGA  | 120 |
| CTTGTTGAGC | GGCTTCTTTT  | GTTGAAATA   | CCGATGTGA  | TGTCACGTGC | GTTTTCGTTG  | 180 |
| TATTGCTCTT | GGAAAAATGAC | AAAGTAATTT  | TCATGCTTGG | CATTTTGTGA | TCTTTTGCAC  | 240 |
| CTTGAGGTGT | CGCAAAATGAG | TCAATATATT  | CAATTAAATA | TGGACGCTTA | ACCGTTTTAT  | 300 |
| ATTCTGCTAA | TGTATAGCTG  | GTCATTGTAG  | GTGTTTTAAT | CGCATAAAAA | GCATCACCCG  | 360 |
| CTGAAACAGC | ATTAAAGCGA  | AACACTTTGG  | TACTAGCGTC | TTTGGAGAGA | AACCATTTTT  | 420 |
| CAAATAATGC | TTTkgTCGTA  | TAGGCATCAA  | ACACCTTTTC | AATCGGTGCT | TCTATCGTTC  | 480 |
| TTGAAAAAAT | AATTTTATTA  | TCTTCAACTT  | TAATCGTCAT | CTTCCCACTC | CCTATCTTTT  | 540 |
| ATATTTTACT | ATCTTAACAT  | GCAATTGAGTA | ACATATGTTA | CCAAATCATT | TGTTATAAAC  | 600 |
| TATAGTAATA | TTAATCATCT  | CCTCTGTnGA  | AATTGAATAC | TACGAAATAA | AAAAACACAT  | 660 |
| CTTATCAGCT | ATTACACGTT  | GAGTAGAGTG  | TGCAAACTCG | TTGAATTAAT | CGCGTAAACAT | 720 |
| TTCTCTTTTG | ATTGGGCTCT  | nATnTAATTT  | TGAAGGATAA | TGTTTGTGCC | AGTTA       | 775 |

(2) INFORMATION FOR SEQ ID NO: 611:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 728 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 611:

ATCAAATTT GATACTAAG GGCTCACGCT ATTTTATGTA TTTATTGGAT TGATTATGTT 60





## EP 0 786 519 A2

TCAATGCATA TTCAACTAGC TCAGGTTTAC TCITTAAGCC AAGCTTTGTC ATAATATGCG 840  
 TCTTATGTGC TTCTACTGTT TTCACAGATA CAAATAATTT CTCTGCAATT TCTTTATTCG 900  
 5 CGTAACCTTT GGC 913

## (2) INFORMATION FOR SEQ ID NO: 613:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 654 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

15

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 613:

ACAAGTGTG ATGTCATAGT AGGTCAAWT GTGCTTTATT TCTGGGACAC ATTAAATTT 60  
 20 TTCCACCCCA TTACTTGGAG CTTAATCTTT ATTACAATTT TACTATTATT AAACATTTTT  
 TCTGTAAAT CATTTGGAGA AACTGAGTTT TGGTTATCAT TGATTAAAGT GTTAACAATT 180  
 ATCGTATTGC TTATTTTGG CTTTTTAATG ATTTTCGGTA TCTTAGGTGG TCATACATAT 240  
 25 GGATTTGAAA ACTATACAAA AGGCCAAGCA CCGTTTGTG GTGGTATCTC TGGTTTCTTA 300  
 GCGTATTAT TAGTCGCCG ATTTTCGGTT GGTGGTACAG AAGTAGTAGC AGTAACTGCT 360  
 30 GGTGAATCAG ATGACCCTAA AAGTCTATG CCTAAGGCAA TTAAACAAGT ATTTTGGCGT 420  
 ATTCTTTTAT TCTATGTCTT ATCAATTGCA GTAATTGGTG CAATTATTCC GTACACAGAT 480  
 CCATCATTAT TAAGAGCAAG TAGTTCAATA AGTCAAAGCC CATTTACAAT TGTATTGAT 540  
 35 AGAGTAGCA TAGCCTTTGC AGCATCAGTA ATCAACGCGG TTATTTTAAC TTCATTATTA 600  
 TCGCTGCAA ATTCAGGTGT TTATACAACA GGCAGAATGT TGTATTCTT AAGT 654

40

## (2) INFORMATION FOR SEQ ID NO: 614:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 642 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

45

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 614:

TCGATCTTTA TTGATAATAA TTAATTGTG GCCTTTAAAA TGTGATATTA ATCCTGCGGC 60  
 50 AGGnTGTA CAAGAGTAGT AACCTAGTAC AACAGGGTG TCAGCATGTT CAATTTTATT 120  
 TAATGCCCTT ATGATGGTAG GTTGATCTAA CATTTCACCG TATAATAGA TGTCCGGTCG 180

55

## EP 0 786 519 A2

5  
 10  
 15  
 20  
 25  
 30  
 35  
 40  
 45  
 50  
 55

ATAAGATTTA TGACATACAT TACAATAAAA ACGATTTAAC GTGCCATGTA ATTCATCAAC 300  
 ATGTTGACTT CCAGCGTCTG AGTGCAAACC ATCGATATTT TGCCTGATGA CACCTAAAGA 360  
 TTGTTGATTA CGTTCTAATT TTGCAATCCA ATCATGAACG ATATTGGGCA TCGTATCGAC 420  
 AAATAGTAAG CGTTTATGGC AGAAATTGAT AAAACCTTCA GGATCATCTT CTAATAATC 480  
 ACGGCTTAAC AAGTATCTG GCGAAAGCCC ATCTTTTGaa ATTTCATCAA ATAAGCCACC 540  
 CATTGaaACGG AAATCTGGAa CGCCACTTGC GACAGATACA CCAGCACCTG TAAAAAATGT 600  
 AATACGATTC GAACATCTTA TAATATGITT TAGTGTCICT AA 642

(2) INFORMATION FOR SEQ ID NO: 615:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1210 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 615:

25  
 30  
 35  
 40  
 45  
 50  
 55

TTTCCCNcC CnCCAAATA TCCAAnGGAA CTTTAATAGT CCAATTGGCA CAGTAAAACT 60  
 ATGGCATTTT ATAAGTATAA TATATCTGTA ATTTATGGTC AATTAGTAAA TTGTTTTTTA 120  
 TTTGAAACAT ATTTACATC AAAATCACAA AGACTTTTAG ATTTGTcTCT AAAAACTCTCT 180  
 TAATAATTTA TTTAATGAGA AGAGTTGCTT ATATAGTAAA TTGTGAAGCC GTTAAACAA 240  
 CGTTACAAA CCTATATCTT TAATACGGAA CCATATGGTA TGAATCAAGG AACTACTTAA 300  
 CTAACAATTC TCTATCAGAT TTATTTGTGT CGAAATCAAC AACTTTAATT GCTTGCCCTT 360  
 CATTTAATGG ATAATTGCT TGCSTAATTT TAACTTTAC AATTGGACCT ATGAGTGATT 420  
 CGTCACCTTC AAATTGTA CTTCATATAAT TATCTGCATA TCCAACTAAT GTACCTTCTG 480  
 TGcCACCCGT TTCTCAGGA ATTACTTCAA GCACATCTTG ATCAAATTTA GACGCATATA 540  
 ACTTTCCGAG TTGATTGCTT AGCGTAATTA ACTTATGCAC CGGTCATTT TTAATTTCTT 600  
 CATCAATTG GTCATCCATT CTTGCAGCTG CGGTGCCAAT TCTAGGAGAA TAAGGGAAAA 660  
 CATGCAGTTC AGAGAACTTA TGCTTTACGA TAAATCATA TGTTCTCTGG AACTCAGCTT 720  
 CAGTTTCACC TGGGAACCA ACAATTACAT CACTCGTAAY TGCCAAGTCT GGTAAAGCTT 780  
 TATGCAATTT TGTTAATCGT TCTGAAAATC TATCCATTGT ATACTTACGT CTCATACGTT 840  
 TTAATCATGT ATCTGAACCA GATTGTAATG GAATATGCAA ATGACGCACA ACTTTTGTG 900  
 AAGCTTCTAA AACGTCAATT ACTTCATCTG TAAGTTGACT TGCTTCAATT GAAGAAATTC 960

EP 0 786 519 A2

|            |            |            |            |            |            |      |
|------------|------------|------------|------------|------------|------------|------|
| CTTTTAAATC | TTGACCATAT | CCACCTGTAT | GAATTCCCGT | CAATACAATT | TCCTTATATC | 1080 |
| CTGAATTCAC | TAGTTGCGTC | GCTTGTTCAA | CTACTTTTTT | CGGATCTCTT | GAAAGCATT  | 1140 |
| AGCCACGAGC | CCATGGAATA | ATACAGAATG | TGCAGAAGTT | GTTACAACCT | TCTTGAATTT | 1200 |
| TTAATGACGC |            |            |            |            |            | 1210 |

(2) INFORMATION FOR SEQ ID NO: 616:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 652 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEO ID NO: 616:

|    |            |            |             |             |             |             |     |
|----|------------|------------|-------------|-------------|-------------|-------------|-----|
| 20 | TAATAAAATA | CAATACTTTT | CAATACAGAG  | AATCCCGCA   | TGTCGGGATT  | CTTATTIATG  | 60  |
|    | CTGATTTTGT | TTTGTCTAT  | GATCAGGGAC  | TTcCAGGSAc  | TCAATTAATT  | ATCACATTCA  | 120 |
| 25 | TATTTCTCTA | TGACCAAAAT | GATACATATG  | ATGCTATTTA  | ATGAGAAAAA  | TAGGCATCAC  | 180 |
|    | TTGGTTATTG | AATTTCTCTC | ATTAAACCTT  | CCAGCTCAAT  | TTAATAGTTA  | GTGCACTATT  | 240 |
|    | ATTCAATAAA | CACCTTTTAA | TCATAAAAAA  | GTGTTTTTga  | TAATTCACTa  | CcAAAAACAC  | 300 |
| 30 | CTTCTTACTT | ATAATTCTAT | TTGTTCAACCA | TTTCTAATCT  | TATCGGCTAA  | ATCATTCAGT  | 360 |
|    | TTTCTTAATC | GGTGATTTAC | ACCTGATTTT  | GAAATTGGAC  | CAGTTGATAC  | CATTTCCTCA  | 420 |
|    | AGCTCTTTCA | ACGAAATTTT | TTGATGTCTT  | ACTCGAATT   | TAGCAATCTC  | TCTCAACCTG  | 480 |
| 35 | TCGTGTAAT  | TTTCAATACC | AATTTCTTTA  | TCaCTAATT   | TAATGCTCTC  | AAC TTGTTTC | 540 |
|    | ATCGCAGCAC | TAAGTCTTTT | ATTTTAGATTG | GcCGTTTTCAC | AAATTAACGAG | TCGGTTAAACA | 600 |
|    | GAATTACGCA | TATCTCTTAC | AATACGTACG  | TCTTCAAATT  | TTAATTAACGC | TT          | 652 |

(2) INFORMATION FOR SEQ ID NO: 617:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 798 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 617:

55

|   |     |
|---|-----|
| TtACCACTTT AAAATAGCGC TTAATAAAATG AAGGGGGCAA GTCTTATGAC GtTTTCaAT | 60  |
| TTCATCaTGG GTTTTCAAAA TGrTAACACA CCATTGGTA TATTGGCCGa ACACGTTAGT  | 120 |

## EP 0 786 519 A2

TCTAATTACA CAGATCATCA ATTAATTGAA ACTACAAATA GAGCTATTAG CTTATATATG 240  
 5 GCAAAATTAAT TTGAGTAGTA CCAATTATGA TGTATTAGTG CATCCCAAT ATCTTTTGT 300  
 TTAAGTTTA TTTCATCATT TCTTATCGAA AATGGTGAA TAATGTCCTT ATCTAACCAA 360  
 GTGTGATAA GTTCATTGG TACACCATCT AACAACTTT CACTTTTACT AATTATAAAA 420  
 10 CATTCOCAGT CAAGTGAAC ATTTTGTGGA TTCACATAAT TACATTGATT ATGATTATCC 480  
 ATAACACTC ACTCCTTTAA AATTCTGTAC TCTTCATTGC GTTTTACCCC GTCACATTAT 540  
 CTTTAAACT AAAATCATCA TTACTTATGA AAAAAATGTA CATCAAAAGC AAAGGTTTT 600  
 15 GCTACCGAAA AAGTTTAAAT AATGGTtAA TATATTtGGT ACTCATTTTA ATAAAAAGAG 660  
 AATACATTTT GAGCTATCAA TACTTTTTAT TGAAGAGGTG TTATTsyTGG CTAAAACGTT 720  
 ATATTTAATG CGCCACGGAC AAACTTTGTt TAATTTTAAG GGACTAATTc AGGGATTtGG 780  
 20 AGATTGCGCG CTAAACAGA 798

(2) INFORMATION FOR SEQ ID NO: 618:

25 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1786 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 618:

ATACGCTCAA TTGATAAAAA TTTAATATCT GCCATTGTGA TAGCATCGC TTCGAATGTT 60  
 35 TCOGTTGCAA TATCAATGA TAATTGTGAC TGGAACCTTA AACATCGAAT CATACGTAAA 120  
 GCATCTCTCT GGAATCGTTC CTCAGTATA CCTACAGTTC TTATTATTG ATTATTAAATA 180  
 40 TCTGTGTGAC CATCAAAATA ATCATACAAT TTGTATGCTG TATCCATTGC TATCGCATTC 240  
 ATCGTGAAT CTGTCGTTG CAAATCTTCG TATAAATCAC GAACAAATGT AACACCACTT 300  
 GGTCACGGT GATCGACATA ATCTTCTTCA GCCCGGAATG TTGTCACTTC ATAATTtCA 360  
 45 TCATtAAAA CTACATTtAT CGTGCCAtGT yCTTnACCTA CAGGTATCGT ATGACTAAAG 420  
 ATAGATTCTA TTTTCATCCG CGTTGCACTT GTTGTGATAT CTATATCATG AATATTCTTT 480  
 CCCATGACAT AATCTCTTAC AGAGCCACCT ACATAATATG CTTCAAAACC ATTGCTTTGA 540  
 50 ATTTGTCTTA ATATAGGCCT TGCCGTGTCA AATAATGATT TATCCATATT ATTACTCGCC 600  
 TTTACTTTTG TTATGCTCAT TTAGCATTTT TTGATAATAA TACTCATATT GATCTGTAAT 660  
 AAGTTCGTGAT CCAAAACGTT CAGCAATATC TGCTAGCATG TTTTCTGAA GTTtGTGTGA 720  
 55

## EP 0 786 519 A2

|    |  |      |
|----|--|------|
|    | ATCTACGACA AATCCAGTTT CACCATGTTT AATAACCTCT TTAATTCAC CGGCATTGTA   | 840  |
| 5  | ACCAATTGGA ACGAGCGCTG TwTTTCATAGC CTCAGTAAA GTTAGTCCA AGCTTTCTTT   | 900  |
|    | TTCACCTAAT AATAATACTA AGTCAGATAA TTGGTAAAA TCACCTACGC AATCITGTTT   | 960  |
|    | CCCTAAAAAT AAAACATCCT CTCTACGTT TAACCTTTC GTCAATTGAC GCATTGGCAC    | 1020 |
| 10 | TAATTCAGGA CCATCTCCAA GTAAAAATA TTTACTAGGT ATCTTTTCAC GTACTTTTGC   | 1080 |
|    | AAATGTTTCT ATAATAGTAT CTATGCGTTT TACTTGCTA AAATTCGATA CATGTATTAA   | 1140 |
|    | CACTTTTTCA TCTGGTGCTA TACCAAAATTG TGATTTTAAT GCTGTGTTAT GTTTAGTTGG | 1200 |
| 15 | AAACTCATT TCACGTACAA AATTATAAAT CGGTATAATT TCITTTGTTAG TTTCAATAAT  | 1260 |
|    | TTCAATGTTT TCITTTGCTA AAGATTACT CACACTTGTC ACAATATCAC TTTTTTCAAT   | 1320 |
|    | GCCAAATTA ATTGCACCTT GGAGTGAATG ATCATAGCCC AAAACAGTAA TATCAGTACC   | 1380 |
| 20 | GTGTAGCGTT GTCATAATT TTATATCTTT ACCTGACATC TCACGAGCTA AAATCCACA    | 1440 |
|    | AATTGCATGA GGTACAGCAT AGTCATATG CAACAAATCA AGATCATATT CTTTAATAAC   | 1500 |
| 25 | TTCAGCGATT TTAGTACTTA ACGTAATATC ATACGGTGA TACTGAAATA CTGCATATTG   | 1560 |
|    | ATTCACTTCA ACTTGATGAA AAATCATATT CGGTAATGGT TTTCTTATTC TAAACGGGAT  | 1620 |
|    | ATTTGAAGTG ATAAAAATGA cTTCTGACC TCGCTCTGCT AATTTAATTC CTAATTCTGT   | 1680 |
| 30 | GGCAATAATT CCAGAACCAC CCATGGACGG GTAACATGTT ATACCTATCh TCATTTCGCTT | 1740 |
|    | GCCCATCCTT TCITTTCTATT TCTChCTATG AthCTCGATG CGTAGA                | 1786 |

## (2) INFORMATION FOR SEQ ID NO: 619:

|    |                               |
|----|-------------------------------|
| 35 | (i) SEQUENCE CHARACTERISTICS: |
|    | (A) LENGTH: 844 base pairs    |
|    | (B) TYPE: nucleic acid        |
|    | (C) STRANDEDNESS: double      |
| 40 | (D) TOPOLOGY: linear          |

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 619:

|    |   |     |
|----|---|-----|
| 45 | ACAGGTATGG ACTThGCTCA AATGACACGA CATTATTAT CAAGACCTAT TGCTATAATC  | 60  |
|    | TTTTGGATCA TTGCAGAACT AGCAATTATC GCTACAGATA TTGCTGAGGT TATTGGTAGT | 120 |
|    | GCTATTGCTC TTAATCTCCT ATTTAACATA CCTTTAATCG TCGGTGCACT AATAACTGTA | 180 |
| 50 | CTTGATGTAT TTTTACTACT TTTTATAATG AAATATGGTT TTAGAAAAAT TGAAGCTATT | 240 |
|    | GTTGGTACAT TAATTTTCAC AGTGTATTTC ATCTTTATAT TTGAAGTCTA TATTTCATCA | 300 |
| 55 | CCACAGTTGA ATGCTGTGTT AAATGGATT ATACCACATA GTGAAATCAT TACAAATAAC  | 360 |

|    |             |             |             |             |             |             |     |
|----|-------------|-------------|-------------|-------------|-------------|-------------|-----|
|    | TTACATTTCAT | CAATTGTACA  | ATCTAGAACA  | TACTCAAGAC  | ATAACAATGA  | AGAAAAAGCG  | 480 |
|    | CAAGGAATTA  | ATTGTGCTACG | ATAGATTCTGA | ACATTTCAGTT | ATCAATCGCA  | TTTGTAAGTCA | 540 |
| 5  | ATTGCTTATT  | ATTAGTGTGA  | GGAGCATCAC  | TATTTTKCAA  | CTCAAAATGCT | GACGATTTAG  | 600 |
|    | GTGGTTTCTA  | TGATTTATAT  | CACGCCTTAA  | AAACTGAACC  | TGTACTAGTG  | GCAACAATGG  | 660 |
| 10 | GTGCAATCAT  | GAGTACATTAA | TTTGCAAGTTG | CATTATTAGC  | GTCAAGGTCAA | AATTCAACGA  | 720 |
|    | TTACTGGTAC  | TTTAGCAGGA  | CAAAATGTAA  | TGGAAGGATT  | TTTAAGATTAA | CACATACCAA  | 780 |
|    | ATTGGTTAAG  | ACGTTTAATT  | ACACGTTCTC  | TTGCTGTATC  | TCCTGTTATC  | GTATGCTTAA  | 840 |
| 15 | TCAT        |             |             |             |             |             | 844 |

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 574 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

|    |            |            |             |             |            |             |     |
|----|------------|------------|-------------|-------------|------------|-------------|-----|
|    | TGTATTGCAG | TCATGCCATA | AATTTTTTGA  | ATAATAAACG  | GAGACGCAGA | AATATAAGTA  | 60  |
| 30 | AATAATATTA | CAAATGTCAT | ACCTTGGAA   | GAGCATTGGT  | AATACAAAAC | GTGGCGCTCT  | 120 |
|    | CANTAATATT | TTGAAGTTTT | TAAACATTGT  | CTTTAAATCA  | CTACTTGACT | CACGATTTGT  | 180 |
|    | CACTGTTAAT | GATTACAGTA | CTTTTAATAA  | AGAAACCTAT  | ACCATGACGA | AGCCAAAGAT  | 240 |
| 35 | AGTCAGAAAT | ACAAAGACCA | TACGCCAGAC  | AGAATAATTT  | AAAATTATGC | CCCTCTATTGT | 300 |
|    | TGGCGCAACA | ACTGGTGCAA | TACCATTAAC  | AAAGCATCAAT | AATGCCATAA | ACTTAGTTAG  | 360 |
|    | TTCATTACCA | CTATACATAT | CACCTTGCAT  | AGCTCTTGAA  | ATAACTGCTG | TCGCCACCTT  | 420 |
| 40 | GTCACCTCTT | GAAGAATCT  | TAAATGCAAC  | ATCAGCCAAA  | TATTATGTAC | AAAAACAATA  | 480 |
|    | CTCTAACTTG | CTAATGTATA | AATAATCAAT  | ATGCTAATAA  | GCGGCTTTCT | GCGCCCAGTT  | 540 |
| 45 | GAATCTGAAA | TTGACACNAG | AACCAAAATTA | CCAA        |            |             | 574 |

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 523 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## EP 0 786 519 A2

|    |   |     |
|----|---|-----|
|    | AAGCAAGAGA GAGTACAACA TTTATATGAT ATTAAGACT TACATCGATA CTACTCATCA  | 60  |
| 5  | GAAAGTTTTG AATTCAGTAA TATTAGTGGT AAGGTTGAAA ATTATAACGG TTCTAACGTT | 120 |
|    | GTACGCTTTA ACCAAGAAAA TCAAATCAC CAATTATTCT TATTAGTAA AGATAAAGAG   | 180 |
|    | AAATATAAAG AAGGCATTGA AGGCAAAGAT GTCCTTGTGG TAAAAGAATT AATTGATCCA | 240 |
| 10 | AACGGTAGAT TATCTACTGT TGGTGGTGTG ACTAAGAAAA ATAACAAATC TTCTGAAACT | 300 |
|    | AATACACATT TATTGTGTTA TAAAGTGTAT GCCGGAATTT TAGATGCATC AATTGACTCA | 360 |
|    | TTTTCAATTA ATAAAGAAGA AGTTTCACTG AAAGAACTTG aTTTCAAAAT TAGACAACAT | 420 |
| 15 | TTAGTTAAAA ATTATGGTTT ATATAAAGGT ACGACTAAAT ACGGTAAGAT CACTATCAAT | 480 |
|    | TTGAAAGATG GAGAAAAGCA AGAAATTGAT TTAGGTGATA AAT                   | 523 |

(2) INFORMATION FOR SEQ ID NO: 622:

|    |                               |
|----|-------------------------------|
|    | (i) SEQUENCE CHARACTERISTICS: |
|    | (A) LENGTH: 1871 base pairs   |
|    | (B) TYPE: nucleic acid        |
|    | (C) STRANDEDNESS: double      |
| 25 | (D) TOPOLOGY: linear          |

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 622:

|    |  |     |
|----|--|-----|
| 30 | AGAAGTGTGA nAAAAATTTA AnAGAGATAT GCACATAGAT GACGCATTGC TATATCCAAG  | 60  |
|    | CAATTGAGAA AGCTGCTGAT GCTCCAAATC aCGGAATGAG GGAACCATGG AGAGTTGTGC  | 120 |
|    | ATGTTCCGAA AGACAGATTA GGAGATATGA GTAAGGATAT TTCTAAATTT GCATTTCCTA  | 180 |
| 35 | ATGAATTAGA TAAGCAACAA TGTCATTATG ATGCAGTTAC GAACTAGGT GGCATGTTAT   | 240 |
|    | TGCTTATTTT AAAACAGAT CCAAGACAAC GTCAAAATGA TGAAACTAC TTTGCATTTG    | 300 |
| 40 | tGCATATGCA CAAATCTTA TGTGTTACT TTATGAAGCG GGAATAGGTA CATGTTGGAA    | 360 |
|    | aTCGCCATTA TATATCTATG ATCTAAAGT AAGnAAACAC TTGGTATAAA GnaAGATGAA   | 420 |
|    | GTTCTTGCTG GATTCTTATA TTTAACGGAT TTAGAAGnAG ATATGCCTAA AGCACCACGT  | 480 |
| 45 | AAAAATAGAA ACITAAATTAC ATTATATTAA TATGTATAAT TATAGAAACA TTAATAAAAG | 540 |
|    | CTGAGTCATG AATTGATGGA CATCTATCGA GTTAGAGATT TAATCTAACT TACTAGAGTC  | 600 |
|    | GGTACAATAA CAGTCTCAGC TTTTATTGTT GCAGTATATA CACATTTTTT TTTTAGTATT  | 660 |
| 50 | TATTtAAAAG TtTCGCTaa AAATGATTCA ACTTGTTCAG GTGACTTAGC ATTTGCTGAA   | 720 |
|    | TGAAGGTGTG CAATTTTATC GCCGTTTTTA AATACTAGCA AGCTAGGGAT ACCCATAACT  | 780 |
| 55 | TCATTTTCAA CAACTACATC TTCTAAATTCA TCACGATTAA CAGTATACCA TTGGTAATCA | 840 |



|    |            |            |            |            |             |            |      |
|----|------------|------------|------------|------------|-------------|------------|------|
|    | CCTGCCTCAA | ATTTAACAAT | TACAGGTGTA | TCGCTATTAA | TTACAGATTT  | AAATGATTCA | 960  |
| 5  | TTACTTTTGA | TTGATTGCAT | TGTAACAACT | CCTCTAGATA | GTTTAATAAT  | TTTTATTATA | 1020 |
|    | GCTAAATTTA | TATCATATA  | AAAAATTTTA | GCTTCAAAAT | GAAAGCCTTT  | TGCTTTGGAA | 1080 |
|    | AATGATATAT | TTATTTTAAA | TACATAAAGG | AGGTTGCAGT | CGTATGATTA  | AATTTTACCA | 1140 |
| 10 | ATATAAGAA  | TGTACAACIT | GTA AAAAGG | AGCAAAGTTT | TTAGATGAAT  | ATGGCGTAA  | 1200 |
|    | TTATGAACCA | ATTGATATCG | TTCAACATAC | ACCTACAATA | AATGAATTTA  | AAACAATAAT | 1260 |
|    | TGCAATACA  | GGCGTAGAAA | TTAATAAATT | GTTTAATACA | CACGGCGCGA  | AATATCGTGA | 1320 |
| 15 | GCTTGATTGG | AAAAATAAAT | TACAAACTTT | ATCAGATGAT | GAAAAGTTAG  | AGTTGTTATC | 1380 |
|    | ATCTGATGGT | ATGTTAGTAA | AGCGTCCTCT | AGCAGTAATG | GGCGATAAGA  | TAACATTAGG | 1440 |
|    | ATTTAAAGAA | GATCAATATA | AAGAGACTTG | GTTAGCGTAA | GTGAATGTGA  | AGCGTTTACT | 1500 |
| 20 | AAATATCTCG | ATATTTAGAT | TCATTACATG | TAAATGAAA  | TAAGCTATAC  | AATTGTTAAT | 1560 |
|    | TTTTATAAAT | ATAGTTGAAT | AGCATCTAGC | CTTAGGCGAT | CATTAAATGAT | GTAAGATTA  | 1620 |
| 25 | ATTAGGAGGG | GATTCCTCTG | GCAGTACCAA | ATGAAYTGAA | ATATTCAAAA  | GAGCATGAAT | 1680 |
|    | GGGTTAAAGT | TGAAGGTAAT | GTAGCAATAA | TTGGAATCAC | AGAATACGCA  | CAAAGCGAGT | 1740 |
|    | TAGGTGATAT | TGTTTTCGTT | GAATTACCAG | AAACAGATGA | TGAAATTAAT  | GAAGGGGATA | 1800 |
| 30 | CGTTGGTAG  | CGTAGAATCA | GTAAAACTG  | TATCAGAATT | ATATGCCACA  | ATCTCTGGTA | 1860 |
|    | AAGTAGTTnA | A          |            |            |             |            | 1871 |

(2) INFORMATION FOR SEQ ID NO: 623:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 451 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 623:

|    |            |            |            |            |            |            |     |
|----|------------|------------|------------|------------|------------|------------|-----|
| 45 | TTGCTGTAGA | AGATAAAGAG | TCCCATCAAT | GGATAGGCTT | TATAGGTTTG | AATTATATTC | 60  |
|    | CAGAAACAAG | CGATTATCCA | TTTAAAGAAT | TACCGCTTTA | TGAAATAGGT | TGGCGCTTGT | 120 |
|    | TGCCAGAATT | TTGGGGAAAA | GGATTAGCAA | CTGAAGGCGC | AAAGGCAACA | TTGAAGTTAG | 180 |
| 50 | CAGAAGAACA | TCAATATAC  | GATGTCTATA | GTTTACAGC  | AGAAGCAAA  | AAAGCTTCAC | 240 |
|    | AACGTGTAAT | GGAAAAAAT  | GGCATGACAG | TGTATGATCA | TTTCGAATTA | CCCAATCTAA | 300 |
| 55 | GTAAGTATCA | TTTATTAAAA | AGGCAAGTGC | GCTATTACAT | TAATCTCCG  | AAAGTGAAAA | 360 |

ATTAGGGGGT GtTnGtKn ATTTTTTAA n

451

(2) INFORMATION FOR SEQ ID NO: 624:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 665 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 624:

|            |            |            |            |            |            |     |
|------------|------------|------------|------------|------------|------------|-----|
| CGATAGTTAT | CTAAAGCCAT | TTTnGATTGT | GTTATGAAAT | CTAATGATGC | GTGATAATTT | 60  |
| AATGCFACAT | AACGrTaATA | TAAAATATCA | ATAGTGAACA | TTTGAGCAGA | TAATGAAGTT | 120 |
| GTTGCTCCCA | TATGCATCTC | ATTTTCATCA | GTTTCCCAT  | AAGTTAAAC  | AATATTGTAT | 180 |
| GCCTGTGCTA | CGGATTATC  | CCTTGACTA  | GTAATTGTAA | TTATAGGTAT | ATGGTAGTCA | 240 |
| TCAATAACTT | TAACCATTGA | TTGCATTICA | CTTTGCGTAC | CATTGTTAGT | AATAAGAATA | 300 |
| aCACTGTGCT | TCGAATTGTG | AGTTGCTAAT | AATGTAGCAA | AAATATGTGT | TTCTTGAACA | 360 |
| AGTTGAATAT | TAAGACCTAT | TCCTGATAAC | TTTTGGTATA | AGTCGGTAGC | AACTACAAA  | 420 |
| GATGCGCCAA | AACCATATAT | AAAAATCGTC | TCAGAACGTT | TTAAACAATG | ACATATTTGA | 480 |
| TCAATAGITT | TATCATTTAA | TTCTTTATTT | GCATGATTAA | GTGCGCGTGT | AGTACGTGTA | 540 |
| TGGAGTTTAG | TTCTTAAAGA | TTCTGTGCTT | TCGTTATTCA | TTAATTCAAC | ATTGTAAATT | 600 |
| GATGATGCTT | TAGGAACATA | TTTAGATATA | TTTATTTTCA | AGTCGTGAAA | ACGCCATCA  | 660 |
| GTAAT      |            |            |            |            |            | 665 |

(2) INFORMATION FOR SEQ ID NO: 625:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 2549 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 625:

|            |            |            |            |            |            |     |
|------------|------------|------------|------------|------------|------------|-----|
| GATGCCAATT | AACCGTGATC | ATAATGTGTA | GAAGTTAATC | GAAGCAATTC | AATATTATCA | 60  |
| AGAAAAAACA | AATCGTCGTG | TTACTTTTGA | ATATGCTCTG | TTTGGTGGTG | TGAATGACCA | 120 |
| ACTAGAACAT | GCAAGAGAAT | TAGCACATTT | AwTAAAGGC  | TTAACTGCC  | ATGTTAACTT | 180 |
| AATTCCtGTC | AACCATGTTC | CAGAAAGAAA | TTATGTGAAA | ACGGCTAAAA | ATGATATCTT | 240 |

## EP 0 786 519 A2

|    |   |      |
|----|---|------|
|    | TTGGATATT GACGAGCTT GTGGTCAATT AAGAGCAAAG GAACGACAAG TAGAAACGAG   | 360  |
|    | GTAAGACAA ATGCTAGAGG CACAATTTTT TACTGATACT GGACAACATA GAGATAAGAA  | 420  |
| 5  | TGAAGATCGG GGTGGTATTT TTTATAATCA AACTAATCAA CAACTTTATG TTCTGTGTGA | 480  |
|    | TGGTATGGGT GGCCATAAAG CAGGAGAAGT TGCAGTAATA TTTGTTACAG ATGAGTTGAA | 540  |
|    | ATCyCGTTTT GAAGCGGAAA ATCTTATAGA ACAACATCAA CGTGAAAATT GGTTCGCTAA | 600  |
| 10 | TAATATAAAA GATATAAATT TTCAGTTATA TCACATGCAA CAAGAAAATG CAGAATATAA | 660  |
|    | AGGTATGGGT ACAACATGTG TTTGTGCACT TGTTTTTGAA AAATCAGTTG TGATAGCAAA | 720  |
| 15 | TGTCGGTGAT TCTAGAGCCT ATGTTATTAA TAGTCGACAA ATTGAACAAA TTACTAGTGA | 780  |
|    | TCACCTATTT GTTAATCATC TTGTTTTAAC GGGTCAAATT ACGCCGGAAG AAGCATTAC  | 840  |
|    | ACATCCACAA CGTAATATTA TTACGAAGGT GATGGGCACA GATAAACGTG TGATCCAGA  | 900  |
| 20 | TTTGTTTATT AAGCGATTAA ATTTTTATGA TTATTTATTA TTAAATTCAG ATGGATTAAC | 960  |
|    | TGATTATGTT AAGACAATG AAATTAAGCG TTTGTTAGTA AAGAAGGTA CAATAGAAGA   | 1020 |
|    | TCATGGTGAT CAATTAAATG AATTGGCAAT AGATAACCAT TCGAAAGATA ACGTTACTTT | 1080 |
| 25 | CATACTCGCG GCTATTGAAG GTGATAAAGT ATGATAGGTA AAATAATAAA TGAACGATAT | 1140 |
|    | AAATTTGAG ATAAGCTTGG CGGCGGTGGC ATGAGTACCG TTTATCTTGC TGAAGATACG  | 1200 |
| 30 | ATACTTAACA TTAAGTTTGC AATTAAGCGC ATTTTATATC CACCTAGAGA AAAAGAAGAA | 1260 |
|    | ACATTAAAAA GTTTTGAACG AGAAGTACAT AACTCATCAC AGCTATCACA TCAAAATATA | 1320 |
|    | GTAAGTATGA TCGATGTTGA TGAAGAAGAT GACTGTTACT ACTTAGTAAT GGAATATATC | 1380 |
| 35 | GAAGGTCGGA CTTTGTCTGA GTATATTGAA AGTCATGGGC CATTAAGTGT TGACACAGCG | 1440 |
|    | ATTAAATTTA CGAATCAAAAT ATTGGATGGC ATTAACATG CGCATGATAT GCGTATTGTA | 1500 |
|    | CATAGAGATA TTAAGCCACA AAATATATTA ATTGACAGCA ATAAAACGTT GAAAATATTT | 1560 |
| 40 | GATTTTGGAA TTGCTAAAGC TTTAAGTGAG ACGTCTTTAA CTCAGACTAA TCATGTGTTA | 1620 |
|    | GGTACTGTGC AGTACTTTTC GCCAGAACAA GCAAAAGGTG AGGCAACGGA TGAATGTACA | 1680 |
|    | GATATTTATT CTATAGGTAT KGTGTTATAT GAAATGCTTG TTGGTGAACC ACCCTTTAAT | 1740 |
| 45 | GGAGAAACTG CAGTTAGCAT TGCGATTAAA CATATTCAGG ATTCGTGACC AAATGTGACA | 1800 |
|    | ACAGATGTAC GTAAGGATAT TCCGCAATCT TTAAGTAATG TCATTTTACG CGCTACAGAA | 1860 |
| 50 | AAGAGCAaAG CGAATCGTTA CAAAACRAAT CAAGAAATGA AAGATGATTT GAGTAGTGTT | 1920 |
|    | TTACATGAAA ATCGAGCGAA TGAAGATGTC TATGAACTCG ATAAAAAGAA AACGATAGCG | 1980 |
| 55 | GTACCTTTGA AAAAAGAAGA TCTAGCAAAG CATATTAGTG AACATAAGTC GAATCAACCT | 2040 |

EP 0 786 519 A2

AGCCAGAAGG TACGGTGTAC GAACCAAAAC CTA AAAAGAA ATCAACACGA AAGATTGTGC 2160  
TCTTATCACT AATCTTTTCG TTGTTAATGA TTGCACTTGT TTCTTTTGTG GCAATGGCAA 2220  
5 TGTTTGGTAA TAAATACGAA GAGACACCTG ATGTAATCGG GAAATCTGTA AAAGAAGCAG 2280  
AGCAAAATATT CAATAAAAAC AACCTGAAAT TGGGTAAAAT TTCTAGAAGT TATAGTGATA 2340  
AATATCCTGA AAATGAAATT ATTAAGACAA CTCCTAATAC TGGTGAACGT GTTGAACGTG 2400  
10 GTGACAGTGT TGATGTTGTT ATATCAAAGG GCCCTGAAAA GGTTAAAAATG CCAAAATGTCA 2460  
TTGGTTTACC TAAGSAGGAA GCCTTGCAGA AATTAAAAATC GTTAGGTCTT AAAGATGTTA 2520  
CGATTGAAAA AGTATATAAT AATCAAGCG 2549

(2) INFORMATION FOR SEQ ID NO: 626:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 2286 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 626:

TGCTTACTTC GCCTTCAATA CGTACTAATT CATGTCCACA ACTTGGACAA TGGGTTGGCA 60  
TATGATATGT GACAGCATCC TCAGGTCTAC GTTCTGGAAT ACTACGTACA ACTTCAGGTA 120  
30 TGATGTCACC TGCTTTTTC ACTACAACAC TATCACCAAT TCGAATATCT CTGTCATGAA 180  
TTAAATCCTC ATTGTGCAAA GATGCTCTTG aTACAGTTGT ACCAGCTACT TTTACTGGTT 240  
35 CTA AAAATAGC AGTAGGTGTG ACTACACCTG TTCGTCCAAT ACTTAATTCA ATATCTAATA 300  
ATTTAGTTAC TACTTCCTCA GCTGGA AATT TATAAGCAAT GGCCCATCTA GGAGATTTTT 360  
GTGTGAATCC CATCTCATCC TGTTGATCTA AATCATTAACT CTTAATAACA ATCCCATCAA 420  
40 TATCATAAAG TAATGACTCT CTTTGGCTTG TCCATTTTTC AATATACTCT AAAACACCAT 480  
CGATATTATT TACACGCGCT CTATTTTAT TCGTTGTAAA ACCTAAITTA TCTAACTCAT 540  
CTAATGCTTC ACTTTGCGAA CGCGCATTGA AATCAGTGAA ATCATTGACA CTATATATAA 600  
45 ATACGCTTAG CTTTCGTTTT GCCGTTAATT TAGAATCTAA CTGTCTTAAT GATCCCGCAG 660  
CAGCGTTTCT TGGATTnGA AATAACTGCT CATCATTTTT TTCTTTTTCT TCATTTAATC 720  
50 GTAAAAATGA ACGTCTCGGC ATATATGCTT CACCACGAAC TTCTACATTT AATGGTCTTT 780  
TCATTTTCAA AGGTATCGCA TGAATTGTTT TTAAATTTTC GGTAATATCT TCACCTGTTG 840  
TTCCATCACC ACGTGTGAAA CCTTGaACGA AGTATCCATC AACATATTTT AATGATACTG 900





## EP 0 786 519 A2

|    |  |      |
|----|--|------|
|    | GCTAATGGTA CTAATACTTT ATCTCCATAT GTGTCAATAA AGTTATAAAA ATAATCATCT  | 300  |
| 5  | GTTTTTGATA CAAATCCAGC ACGCTCTTCA GTTTCACGAT ATAAATCTAA GAAAGAGATTG | 360  |
|    | AACATCATCAC GTTCAAGGAA TCTGACTTTA ACACCATAGT TTATCGCTTT ATTAATATTA | 420  |
|    | CGTTTACGTT GACIATCAAA TGTCTTTTTC AATGTTTCGG GTGTTTAC CCATAGGTTT    | 480  |
| 10 | AATACGCCCA TCCATCGTAC TTGGCTCGAT GTATCATACT CAGTTGTAAA GCCATGATGC  | 540  |
|    | TCGTAAACAT GTGATTTAAA CAAGTTTACT AGGGCATCAT TTTTCTCGCG ACCTTCAAAT  | 600  |
|    | GGCAGCATAT CTTTATCATA TAGATGATAT AACCAATACG GATCTAATTT AACATATAAA  | 660  |
| 15 | CATTGATGTT GCTGTAAATA TTTATCTAAC TCTTTTAAAT AATAATCAAC TAATCTTAAA  | 720  |
|    | TCTGAAAAAT CCATTACTGG ACCACGATTC GAATAGTAAA CATAACTTCC CATAGTAGGA  | 780  |
|    | ATTTIAGAGA AAAGGCTTGC TGCATTACT TTGTTATTGT CGTCTTAAAT ACCTAATAAA   | 840  |
| 20 | ACTACTTCAA AGCCATCAIT CTCACGGGTA ACTATATTTT CTTTACTTGG GAAATAATGA  | 900  |
|    | CTTTCCAATG ATGGATTTTG TACAAAGTGT TCAAATTCGG TAACAGTTAA CTCTGTAAAT  | 960  |
|    | TTCATGTTT GATAATTCCT TCCTAAAAAA TTCTGCTCTT AACTTTTTTA AGTGGCGTAT   | 1020 |
| 25 | ATGCTGGGTA AACAGGTTTA TTAATTGGTT TAATAAAGTC ACCAACATAT TCMATAATTT  | 1080 |
|    | CAGCATGTGA ACCTTTTTTG AATTAACTA CACCAGCATC TTCAGCATCT TCTGTAAATT   | 1140 |
| 30 | TrCCAATAAC ACCATAGAAA TTATAACGGT CAATGCATGA TTAAATGCAT AATTAATCAT  | 1200 |
|    | TTCCCATTGC ACTGCATAAC T  | 1221 |

(2) INFORMATION FOR SEQ ID NO: 630:

|    |                               |
|----|-------------------------------|
| 35 | (i) SEQUENCE CHARACTERISTICS: |
|    | (A) LENGTH: 1121 base pairs   |
|    | (B) TYPE: nucleic acid        |
|    | (C) STRANDEDNESS: double      |
| 40 | (D) TOPOLOGY: linear          |

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 630:

|    |   |     |
|----|---|-----|
|    | TGGCCCAaNT AACACAAGTA ATGCTGCTA ATGGTAACAT CAAAAAAAT GAAATCGTAA   | 60  |
| 45 | CTACAAATGT TAAACCTTGG AATACACCAA CCATTCTCTG TAATCGTTTA CTATAGTATC | 120 |
|    | TATTGTGAAT CCAAGTAATA ATAGCTGAAA TAATAATACC ACCTAGAATA TTGCTATCCA | 180 |
| 50 | ATGTGGCAAT ACCTGCAATT GATTTTAAAC CAGGTACATT TTCAACGCCT TTTTCTAAAT | 240 |
|    | TAGCGCCAAA CGTATGTGGC CATTGTGTTA AAATGGCATT TATAAATGTA TTAACATTA  | 300 |
|    | AGTAACCCAT CAATGCTGCA AGTCTGTCAT GACCTGGTGC TTTTGTAGCT AAAGAAAGTG | 360 |

55

EP 0 786 519 A2

TTACTGACCA AAATTTAAAC CAAAACGTAT GTTGATCTGC TAAACTCCCC ATGATGTAG 480  
 GATTTTAA TAATGTGCA AAGCCAAGCA CTATCCCAA GAAAGCGAAC ATTAATACCG 540  
 5 GTACAATCAT TGCACACCG AAGCGCTTA TCGCATTAT CTTCATTCC CTCCATATCA 600  
 TCTTCCCTAA CAATACATCT AATTAGATTC ATTTATAAT AGATGCTTA CTATTAAAT 660  
 10 ATAATATATA GTAAACGCTT ACACACCTAC AACGACATTG ACGTATTTG AAGTATTTT 720  
 GTATAATCAG ATTATCTTT CATATAGTGA AAATTTTTC ACGACCTTAT ATATGACATC 780  
 GTTGATTG TAATACATTC GTTTAAACG CATATACAA CCTATATCAA TACACAAATA 840  
 15 TATATAATGA CATACAAGAT TTTAATGTAA TAACGATCTA TTACACATTT ATTTCAAGG 900  
 AGGTGGAATA TGTTTTAGA TGAACACATT AATCGAACT TTGATAAACT TAATGATAAT 960  
 GATTTCATA TCGCTCACTT TATCAATACA CATATAGATG AATGTAAAAA TATGAAAAA 1020  
 20 CAAGATTAG CGCAATTCAC ACATGCCTCG AATGCAACCA TTCACAGATT TACACGCAAA 1080  
 TnAGGTTTGG ACGGTnATAG TGGATTAAAA TCGTACCTTA A 1121

(2) INFORMATION FOR SEQ ID NO: 631:

25 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 4005 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 30 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 631:

35 AACCTTCCAT TTTACTTGAT CGATAACATC AGTTCGTCTT TACAAATCTC TTCATTAATA 60  
 TCGCTCTTAA AACCATGAAA TTTAATCATG TCCGATAAAT GATAATCTTC TACAAGTTGT 120  
 CGATATTCTG ACAAACCAT TCCATGTCCA TAAATATTCA ATTGAATATT GGGATGTTTT 180  
 40 GTTACTAATT GCTTGATTAC TTCAATTGA TGTTTAATTT GTTTATTTTC AACGAGGCGA 240  
 GCAATTGATA TGATATGATT TTTCTCCTTT TGATTGATGT CAAATGATA CTTTAAATTT 300  
 GCCACGTAGC CAACCGGAAT ATTGATAACT GGTATTTTAT TTTCAATATA TTGTGAAATA 360  
 45 TCTTGCGATT GCTTTTCTGT TGATACAACA ATCGCTTTAT AACGTGTAA ATTATTAAAC 420  
 ACTGTTTTAT AAAAATTTT TATACCAITA CCGGCACCGG ATAAATGTGT ACTATGGAGC 480  
 50 ACAACAATAA CTGGAATACT TTGATTTAAT CCGCTATAA CAITTCCTAA TTCATGAGGA 540  
 CGATCTAATA TGATTTGATC ATTATTTTIA CATAATTGAT GSAGAAAAA TTGAACATAA 600  
 TCATCTCTCG TATCAAAAAA TTGTTGATGC TGGTCTTCAT TTAAGATAAC CTTTGTGAGC 660

55



EP 0 786 519 A2

|    |            |            |            |             |             |             |      |
|----|------------|------------|------------|-------------|-------------|-------------|------|
|    | TAGTAATTTT | CGAGTACAAT | CGTTGTCCT  | TCACCTAAAA  | TTCGAGAACA  | ACTTAAAAAG  | 780  |
|    | CCTCTCCAT  | CATACAATC  | GCCTTTTACT | TTTCTCTTT   | TATGATCAAA  | ATAATTACACA | 840  |
| 5  | TAATTTAATT | GATGATACTG | TTTATCTAAA | AAATGAGCAT  | ACATTACAAA  | TTGCTCTTCA  | 900  |
|    | TCATATATTC | TGACATCATT | TGAATTTTCC | ACAAATTTCA  | ATGTGTACCT  | ACATGACTTT  | 960  |
| 10 | TCCCAATACT | GTATCCAGTT | AACCTGCTTT | GTCTTTTTAT  | AATTGATTGC  | TTTTTGAAAA  | 1020 |
|    | TAGTCATACA | TTGTAAATAC | ATCATTTTCA | ATCTGATGTT  | GCTTCGCATA  | TGTGTATGAA  | 1080 |
|    | TAAGGATTC  | ATTTAACATA | TACACATTTT | GAAGATATGC  | CGTGTGTTT   | GAACAACTTC  | 1140 |
| 15 | AATCTAATTA | TTTGCCTTT  | TTCTACACCT | GTAAATTTTAC | TTTCTAAAA   | TGTTCTTAAA  | 1200 |
|    | ATGTAATTC  | TATTATCGCC | TCATATAAGT | TTTATTCGGT  | ATCTTTATTG  | TTTATTTTAT  | 1260 |
|    | ATGAAAAATA | CATCTATTGC | ATGTGTAATT | ATAAAAAAAC  | CAGGCCACAA  | GGACCTGGGT  | 1320 |
| 20 | CATATTGTAT | TATTTGTTTT | GTTTTTTGCG | ACGACCGAAT  | AACAATAATG  | AACCTAATGC  | 1380 |
|    | TGCAAAATAT | CCACCAATA  | ACGTTGCGTT | ATTTGAGCGG  | TTATTTTTCAC | TACCTGTTTC  | 1440 |
|    | TGGTAATGCT | TTTGCTTTAT | TGTGATGGTC | TTTAGTAGTA  | CTCATGGGTT  | TAACAGGTGT  | 1500 |
| 25 | ATGTTTTTCT | GCATCCGAGT | CTGAATCGCT | GTCTGAATCA  | CTGCTGAGT   | CTGAGTCGCT  | 1560 |
|    | ATCAGAGTCT | GAGTCGCTGT | CCGAATCTGA | GTCTGATCT   | GAGTCGAGT   | CGCTGTCTGA  | 1620 |
| 30 | ATCTGAATCA | CTGCTGAGT  | CTGAGTCGCT | ATCTGAGTCT  | GAATCGCTGT  | CTGAATCTGA  | 1680 |
|    | GTCTGATCT  | GAGTCGAAT  | CGCTGTCTGA | ATCTGAGTCG  | CTATCTGAGT  | CCGAATCGCT  | 1740 |
|    | ATCTGAATCT | GAGTCGCTGT | CTGAGTCGTA | GTCTGATCT   | GAGTCGAAT   | CGCTGTCTGA  | 1800 |
| 35 | GTCTGAATCT | CTATCTGAAT | CTGAGTCGCT | GTCTGAGTCT  | GAATCGCTAT  | CTGAATCTGA  | 1860 |
|    | GTCTGATCT  | GAGTCGAAT  | CGCTGTCTGA | ATCTGAGTCG  | CTATCTGAAT  | CTGAGTCGCT  | 1920 |
|    | GTCTGAATCT | GAATCACTGT | CTGAGTCGTA | GTCTGATCT   | GAGTCGAGT   | CGCTGTCTGA  | 1980 |
| 40 | GTCACTATCT | GAGTCGAAT  | CGCTGTCTGA | TGTATCTTCT  | TGGAAGTATC  | CGTTATCAAG  | 2040 |
|    | TGTGAATCA  | TCATGATCCG | TAATTTGTAC | GTCAACTTCG  | CCACCATCTG  | CATCTTTTATC | 2100 |
|    | ATCTTCAGTT | GTATTTGTAA | CTGTTTGTGT | TAAGCCAGCA  | GGCTTTTCAA  | AAATAACTTT  | 2160 |
| 45 | GTATTTACCG | CTATCTAAAT | TATCAAAGCA | GTATTTACCA  | TTTTCATCTG  | TTTATGTTGT  | 2220 |
|    | TCCAATTACT | TGCGCTTTT  | CATTTAATAA | AGTAACCTTA  | ACATCTTTGA  | TACCTTTTTC  | 2280 |
| 50 | AGTTGAATCT | TGTTTGCCGT | CTTTATTACT | GTCTGACCAA  | ACATAATCAC  | CTAAACTATA  | 2340 |
|    | TTTTGGTGT  | TTATAGAAAC | CAGTGTCTAA | TGTCATGTTA  | TCTGCATCTT  | TAATGACACC  | 2400 |
| 55 | TGTTGTGTGT | AAACCATTAG | AATCTTTTTC | AGTATCATTT  | CCAGAAGTTA  | CTGAAGTTGG  | 2460 |

EP 0 786 519 A2

|    |  |      |
|----|--|------|
|    | TTGATATTTA CCATTTTCAT CTGTGTAAAC TGTTTTAAAC ACTTGTGCGT TTTTCATCTT  | 2580 |
|    | TAACGTAAC GTTACACCTG AATGCCCTT TTCATCTTAA TCTTGAACAC CGTTTTTATT    | 2640 |
| 5  | TGTATCTTCC CATACATAGT CACCTAAGTT GTAAGTCGGT TTGTAGAAAC CAGAGTCAAT  | 2700 |
|    | AGTATCGTTA TCTTTATCTT TAATGACACC TGTGTGTGAT GTACCATTTG AATCTATACC  | 2760 |
|    | TTTCATCAGT CCGTAACCTA CTGTGTTGG TGTGTAACT GATGGTGTTC CGAATCAAC     | 2820 |
| 10 | TTTATAAGTT CCATTTTCTA ATCCAGTAAA TTGATATTTA CCATCTTTAT CTGTTTTAGT  | 2880 |
|    | TGTTTGTAAA ACCTCACCGT TTTCAATTTT CAATGTAAC GTTACGCCGT AAATACCTTT   | 2940 |
| 15 | TTCAAGTGAA TCTGCTTAC CATCTTTATT TGTATCTCC CATACATAAT TACCTAAATT    | 3000 |
|    | ATATTTTGGT GTTTTGTAGA ATCCACTATC TAATGTCATG TTATCAGCAC CATTAATAAC  | 3060 |
|    | ACCTGTGTGT GTTAAACCAT TAGAGTCTTT TTCAATGTCG CTACCAGATG TTACTGTAGT  | 3120 |
| 20 | CGGTGTATAG CTTTCTGGTG TAGTAAATTC AACTTTTATA TTACCATTTT CTAATCAGT   | 3180 |
|    | AAATTATAT TTGCCATCAG CGTCTGTTGT AACTGTTTTT AaCagTTACC GTTTTCATCT   | 3240 |
|    | TTTAATGTTA CCGTTACGCC AGATATACCT TTTTCATCTT GGTCTGGAT ACCATTTTTA   | 3300 |
| 25 | TTGTATCTTT CCCAGACATA GTCACCTAAG TTGTATTAG GTTTGTAAAT ACCTAAGTCT   | 3360 |
|    | GCAGATAAGT TATCTTGGCC ATTAAGTGTG ATAAGTGAAG ATAAGCCGTT TGAATCTAAT  | 3420 |
|    | TCTTCGTTAT TACCTTGTTT TGAAGGGGTT ACTTCATAAC CTTTGGTAA GTTTGAAAA    | 3480 |
| 30 | TCTACACGGT AATCTCCATT AGGTAAGTTT GGAATCAAGT ATGACCCATC TTCTTTAGTA  | 3540 |
|    | ACTGCTTCTC CTACTTTTGT ATTTGTATTA TTATCAAAATA CAGTTACAGT TACATTGCCA | 3600 |
| 35 | ACGCCTTTTT CTCCTAATTC TTGAACACCG TTTTATTAG TATCTTCCA TACGTAGTTA    | 3660 |
|    | CCAATTTTAT ATACTTCTTG ACCAGctCCG CCACCTTGGT TATTAGTAAA TCCTAAAGCA  | 3720 |
|    | TTGCCAGTAG AAACGGATTT ATTACCTGTT GAAGATAAAG TAGCCATTGT AACAAAGTGT  | 3780 |
| 40 | GGGCTTTCGC TATTTGTATA TTGGAATTTT GTATTAAACA TTACAACATA AGCAGAATCT  | 3840 |
|    | GCATTTCCAA AATCAATAAC AGCGCTATTG TtGTCGCCAT ATGTAATTTT CTGCAAGTAT  | 3900 |
|    | TGATTGTGTA CATCTGTAAG CTCTTTAGTA TTCACATCGT ATCCTTTTAT TAATGTATAA  | 3960 |
| 45 | CCTTTAGGAA CTGATATAT TTTTATATCT GTTACATCTT TATTT                   | 4005 |

(2) INFORMATION FOR SEQ ID NO: 632:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1440 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

## EP 0 786 519 A2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 632:

|    |  |      |
|----|--|------|
|    | TATTAGTTTA CTCTAGTTTC CAAAGCGGGA ATTTAATGT TATTAACAGC AAGGACAGCA   | 60   |
| 5  | AAAGCAATAT CGGCGCATTG ATTGAAAATC CAGGAATATA TCCTTTTATG TCTGGATATG  | 120  |
|    | AAAACCTGAA GTTATTGAAT GAATCAAAAA ACACTCAAGA TATCGATAAA AITGTCTCAC  | 180  |
| 10 | AACCTCATAT GGATGAATAC ATTCATAAAA AAGCTAAAAC GTATTCTCTT GGTATTGAAAC | 240  |
|    | AAAAATTAGG AATTGCTATA GCATTTTAA ATAAACCTCA ATTCATTATC TTAGATGAAC   | 300  |
|    | CAATGAATGG CTTAGATCCA AAAGCTGTGC GAGATGTACG TGAATTGATT GTCCAAAAAG  | 360  |
| 15 | CGCAAGAAGG TGTACTTTTC TTAATTTTCA GTCATATTTT AAGTGAATTA GTTAAAAATCA | 420  |
|    | CAAACTCTAT CCTATTATT AACAAAGGTA AAATTGTTAC AGAAACATCG GAAGAAGAAC   | 480  |
|    | TTAAACAATT TAAAGATAAT GATTTAGAAA ATGTATTACT AGAAATCATA GAAAGGGAGG  | 540  |
| 20 | ACCAAGCATA AATGGGAAC TTTAATTAAA CAAGAATGTT TCAAATTATT TAAAAAGAAA   | 600  |
|    | TCAACTTTTA TCGCACCTAT TGTCTTTATT CTACTAATGG TTGCTCAAGG TTATATTGCT  | 660  |
|    | ACAAAATACA ATGAAATTTT TACGCCACAG GAATCTTTCA CATCTGCTTA TAATGGTTTT  | 720  |
| 25 | TCATGGTTTG CATTTTATT AATTATCAAA GCAAGTACAA TCATTTCAAT GGAATTTTCA   | 780  |
|    | TACGGTACGA TTA AAAATTT ACTCTATCGT GAATATTCAA GAACAACTAT GATTGTTAGC | 840  |
| 30 | AAAATCATCA CATTATTIAT TATTCTTTTA ATTTATTTTG TTATTACAAT TATTGCTTCA  | 900  |
|    | ATTGTTIATG GGTCTTTATT CTTTAATGAT TTAATATAT TTGAAAGTAG CGGTAATCAA   | 960  |
|    | TTATCTTIAT TGAATCAATT ATTATTAGTT AGTTIAGGCA CATTGTTGG CGTTTGGTTA   | 1020 |
| 35 | GTTTTAAGCT TAACGTTGCT ATTATCATCT GCAACAAAT CAACGGGAGT AGCCATTGCT   | 1080 |
|    | GTAGGTATTG TTTTATTATT TGCAAGTTCT ATTTTAGCAG TTATTCAAAC GGCACTTTTA  | 1140 |
|    | GAAAAAATAG ACTGGCTAAA GTGGAATCCT ATTAATATGA TGAATATTAT GCTTCAAAAA  | 1200 |
| 40 | GTGAAAAAG GCTTTAGTAA GTCGACAAAA TTAGAACTTC ATGAATTGTT TATTGGTAAT   | 1260 |
|    | ATTGCTTATA TTCTATTATT CTTAATACTT GTAGTATTTA TTTTCAAGAA GAAAAATATT  | 1320 |
|    | TAGTAACCTA AAGTATTAAA TGCTATAATA CACACATATT CCATCGTAAT TCAAAATCAT  | 1380 |
| 45 | TTTCAAATCC CTTACCCCAA ATAATGGTGC GGGGATTTTT TCATCCAAAT TTTGGAAATT  | 1440 |

(2) INFORMATION FOR SEQ ID NO: 633:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1323 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## EP 0 786 519 A2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 633:

```

GCTGACATAA TTGCATCAA TTTCACATCC CCATAAAAAAT CGCTACCACA TAACCTACGA      60
5  TAATACCTAC AAGAACTGGA ATTAAGATA GGAATCCTTT AAAAAATCCT TGAACGACTA      120
    TTGTTACAAG CAAGGTTATC ATTGCAACAA TTAAGAACT GATATTGTAA CCTTTCATAT      180
    CTCAGGATT TTCATACATT GCCATATTGA CTGCAGTAGG CGCTAAGCTT AAACCAATTA      240
10  CCATGATGAC TGGTCCACAA ACAACTGGTG GTAATAATT CATTAACTAT GCTGTCCAC      300
    TTAATTGAT TAGAATCCCG ATGATGACGT ACATAACACC ACTCATGAAT AATGCTACAA      360
15  GCATGTCTCC TAAGCTATGC GTACTTAATC CCGTGATAAT TGGCGTGATA AATGCAAAGC      420
    TAGATCCCAA GTATGCTGGT ATTGCGCCT TCGTTATTAA GATATAAAGT AATGACCGA      480
    TTCCCGAAGC TAGTAACGCT GCTGATATTG GTAGTCCTGT TAAGAATGGT ACTAGTACTG      540
20  TTGCGCCAAA CATCGCAAT AAATGTTGTA AGCTTAAAAA TGCCCATTCG GCTGGTTGTG      600
    GTTTTTCATT TACATCTAGT ACGGGTTTTA CTGTTGTTTC AAACATTTC AATTTTGTCA      660
    TAATATTCAAT TTCTCCGAT AATAAAAAAA TCTCTTTACA TCAGTATATG TAAAGAGACA      720
25  AAAAGTGTA CAAGTTGCTA CAAGTCATT TCGTCCATAG AAATTGACTT ATAGTTGTG      780
    AACATGAGGG TATTATTAGA TAAACAAGCA TATGAAACT TATTATCAT TCAACTCCCC      840
    CACCTTTTTC AGTCTCTCGT ACTGAATTAA AAGGGGtATT ATTTAATTAT AACTGCATT      900
30  CTTTGATCca TTtCTCyAA ATAGACACTT ACCGTTTCCT CTTTAGAAGT AGGTAWATT      960
    TTACCAACAA AATCTGCTCG AATTGGTAAC TCACGATGTC CTCGATCAAC CAAAGCAGCT      1020
    AAACCAATTT TAATAGGTCT AGCATTTAGC AAAATAGCAT CAAGTGAAGC ACGAACCGTT      1080
35  CGACCAGTAT ACAGCACATC GTCAATAATG ATGACTACTT TATCTGTAAT ATCTGTGTG      1140
    ATGTCTATTG CGTCTTTTGT CGTAAGTGAT GACATGTGct CTATATCATC TCTAAAGTAT      1200
40  GTAATATCAA TTGTCCAGT AGGTATACGT TGTGTCTCAA TTTGATGAAT TTtATCTTGT      1260
    ATACGATTG CTAAATATTC ACCTCTTGTh TTGATACCTA AAAGATTAAA TTATCAGTAC      1320
    CTT
1323

```

(2) INFORMATION FOR SEQ ID NO: 634:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 761 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

|  |    |            |            |            |             |            |            |     |
|--|----|------------|------------|------------|-------------|------------|------------|-----|
|  | 5  | AACCATCTCA | TGTAATTTTG | AATTTAATAG | ACGAGATCGA  | TGTATATCTG | AATTGATCCC | 60  |
|  |    | AACATAGGAT | TAATGTGTGT | ACATCATLAA | cGCATAACCA  | ACATTTTGAG | CAGTGTGTTT | 120 |
|  |    | ATAAGTAACG | TGATTTTTAT | CTAATTGCCC | TCCTAATGCG  | TCCCTGTGTA | ATTCAACACT | 180 |
|  |    | ATCAGAACCA | TTAGAGGTCG | CTTCATATAA | GTTATTAGAT  | GCAATTGTGT | CTAAATCGCT | 240 |
|  | 10 | ATGGAATTTT | AATGGTTTTA | ATCCCTTTTA | TTTTCTCATAT | TCATTCTGTA | CTTCATAAAG | 300 |
|  |    | AGAAATTAAT | TGATTTGGAT | TTTGCTCAAC | TGACGCGTTA  | TTATGCTCTT | CTGACGTAGA | 360 |
|  |    | ATTAGAATT  | AATTGATAAG | GTTCAATATC | TGCTAACATT  | TCTTTTGTTA | AAAATCGTAC | 420 |
|  | 15 | ACTTAGCACC | TTTTTCGATT | GTTGATCAGA | ATACACTTGT  | GCATATATGT | CGCCATATTT | 480 |
|  |    | AATCAGGTGT | TGTGTTTTTA | AATCTTCATC | TGAAAGTTCA  | AATTGATATT | TTTTACCATC | 540 |
|  |    | AACCTTAAAG | GACGGTCTGT | GATTAAATCT | TGTATGATTA  | AAAAATTTCT | GAGAATGTTG | 600 |
|  | 20 | TCCTATTTTT | AACGACTATA | CATTGACTTT | CTTCCTGTGA  | GTAAACACTG | AAACGATTTT | 660 |
|  |    | TTCCAGCTTA | GTTGAAACCA | TGTAATTAAC | GTTCTTGCTC  | CTTAAACGAT | AATTTTTGTA | 720 |
|  |    | rCCATCTCTA | AAAGGCTAGA | CrCGACTCAG | TTGTCCAAAT  | T          |            | 761 |

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 827 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

|             |            |             |            |            |            |     |
|-------------|------------|-------------|------------|------------|------------|-----|
| NTAAATATAT  | TTATATATTA | TAGAATAGAA  | AGACCTGAAG | ATTGAATATC | TTTCGCAAG  | 60  |
| CCTTTAACTG  | TATCTACTGA | TAAATCGTTA  | ATATCGCGAC | CIAAGTTTGT | ATTCACTTTT | 120 |
| TTCAACAACAT | CTGCTGGGCA | TGTAATAATA  | TCTGCACAA  | TTTCATCAGT | TGAATACATA | 180 |
| TTGAATAAAT  | CGCGGCAACT | TGCCCATAAAT | AATTTAAACG | CGTCTTACT  | ATCGGTAACT | 240 |
| TTnACAGCCT  | CTkTCATTA  | TGGTaaATGA  | TCTACGCGTG | TAlCTGCAAT | ACGTCTGTCA | 300 |
| AATACTGAAA  | CATATGTTGG | CACACCTTCA  | GTTACTGCTT | CAGTTATTTC | TTTAACTTGT | 360 |
| TCAATTGTTG  | AAACAGCCGT | AACGTTTAAT  | CTCACATTGT | CAGCTGAAAG | TTTTTTAATT | 420 |
| AAAGGAATCG  | TGATATTACC | TTTTGTATTT  | ACAATAGGAA | TTTTAAACAA | TACATTTACT | 480 |
| CCAAATTTGT  | TTAAAATTGC | TGCTTCTTTT  | TCCATAGTTT | CTAAATTGCT | TCGAAATAGC | 540 |
| CTCAATGATA  | TTGAAGCATC | TGGAAATTTT  | CTGCACGTTT | CTTCAGCAAA | AGGTTTGTA  | 600 |

TTTTHATAAG CTGCTTTCAT ThCTTCAATA TCTGCACCGK CGCAAATAC TTCTACATTT 720  
 AGTTTAGCCA TATAAATAG CCTCCTTGAT TCTTATTAAA ATTITTAACAA CATCTGCATG 780  
 KCTTTTTCTT ACAACCATTT GTAAAAAATG ATTTTTATTT CTTTGTT 827

(2) INFORMATION FOR SEQ ID NO: 636:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1478 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 636:

TTAGTCGGTA TAACCATCGG CTAGGTGGTT TTGTATTAA AAAATGGAT ACCAAAAATTT 60  
 ATTAATAAAT ATTTTAATGT TAGAAAAAAA CTAATAAAA ACTCGCTAAT GATATCCAAT 120  
 AATATGTATA CAAAACGAGA CATATATTGC ATATGATTAA CGAGATACTG AAAATATTTT 180  
 ATCACCCCTA AAATGATTAT TCATTTTCAG CGGTAATTCG ACCTAAAGTC AAACCTACAA 240  
 TAAACCGAT GATAAATACT ACTAATGAAA CGAACCACAT CACGATATTA GTTGSTAAAC 300  
 CTGGAAATAC TGCAAAGAGG GAGCCAACAA CAAAACCAAT GATTAATGCA AAAGTCATTA 360  
 GTTTATGATG TGTAGGAAA TACTGGATAA TTTTGCTTGA AATAATGAAT CCAGCAAGCA 420  
 CGCCAAATCC GACTGCAAGT AATATAGGAA GACCTGCAAA GTTAAGTTTA ACAACTTCAG 480  
 ATATTGCTAG CATGACCGTA CCATAGACGC CAAATACTAA TAACATAAAT GACCTGAAA 540  
 TACCTGGGAG TAACATAGCA CTAGATGCAC ACATACCTGC AATAAAATAT TTAATAATA 600  
 GACTAGTTGA TAGAGTAAGT GTTCTCCAG CATGTTTATC ACCATTATTC ATTAATGTAA 660  
 TAACAATTA GATAGCGATA CCAGCTATAA CCATCATGTA ATGTYTAGTT GTAAATGAAG 720  
 TTTTATAGTT AGAAATTTTC AATAAATATG GAACGATACC AATGATTAAT CCACCAAGA 780  
 AAAACATAGT TGGAAATATG TGTGGCGTTA ATAAATAATT AAAAAGATTA CTTAGTGATC 840  
 CCATTGCCAG TAACATTCCA ATTATAATGG GGATTAAAAA TGTAAAACTT GGCCAAAAAC 900  
 GTCGTAGAAA TATGCCGCTA ATTGAAGCGA TAAATTGATT GTAAATACCT AACAATAATG 960  
 CGATAGTCCC ACCGCTAACA CCAGGTACCA AGTCACTCGT TCCCATAGCA AAACCTTTTA 1020  
 GAATATTAAT CCAATTAAAT TGTGTCATGA ATAACTCCTT TCAAACGATT GGAATAAAAT 1080  
 CATAAATAGC ATCATACCAT ATTACAAATG TCCTAGTGAA ATGATAACAT ATTTTAAAT 1140  
 CATAAAATCC ATTGAGAAAT TATGTGCACT TATTATCATT TATATTTTTA AAGAGAGCGG 1200

## EP 0 786 519 A2

AGGTATAAGT AAGTTATAAT TAACTGAACG CATTATTACA AAGTCITTTT GACTACAAAT 1320  
 TAAAAATTATT ATAACTAGT TAAGAAAACT TTATATTTTA CGGAGGGAAT ATAAATGGC 1380  
 5 ATCAACATTA GAAATYAAAG ACCTACATGT GTCTATTGAG GATAAAGAAA TCTTAAAAAGG 1440  
 TGTTAACCTG ACAATTAACA CTGATGAAAT ACATGCGA 1478  
 (2) INFORMATION FOR SEQ ID NO: 637:  
 10 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1995 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 15 (D) TOPOLOGY: linear  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 637:  
 20 ATTACAGCAT CTTCTCTAGG TAGATTATTA AAAGATAGAG GTCTAAATGT AACAAATCAA 60  
 AAAATCGATC CATACTTAAA TGTGACCCA GGTACAATGA GTCCTTATCA ACATGGTGAA 120  
 GTATTCTGTA nGGATGATGG TGCAGAACT GACCTAGACT TAGGACATTA CGAAAGATTT 180  
 25 ATTGATATTA ATTTAAACAA GTTTTCAAAT GTGACAGCCG GTAAAGTGTA TTCACACGTA 240  
 TTGAAAAAAG AACGTCGTGG TGATTACTTA GCGGGAACAG TTCAAGTTAT TCCGCATATT 300  
 ACAATGAAA TTAAGAAACG TTTATTACTT GCAGGGGAAA GTACGAATGC AGACGTTGTT 360  
 30 ATCACTGAAA TTGGCGGTAC AACAGGTGAT ATTGAGTCAT TACCGTTTAT TGAAGCGATT 420  
 CGTCAAAATC GTAGCGATTT AGGTAGAGAA AATGTTATGT ATGTTCACTG TACATTACTG 480  
 CCTTATATTA AAGCTGCTGG AGAATGAAA ACGAAGCCAA CACAACATAG TGTTAAAGAA 540  
 35 TTACGAGGCT TAGGTATTCA ACCAGACTTA ATCGTTGTAA GAAGTGAATA TGAATGACA 600  
 CAAGATTTAA AAGATAAAAT TGCATTATTC TGTGACATTA ATAAAGAAAG TGTATTGAA 660  
 40 TGTCTGTATG CAGACTCTTT ATACGAAATT CCATTACAAT TAAGCCAACA AAATATGGAT 720  
 GATATCGTTA TTAACGTTT ACAATTAAAC GCGAAATATG AAACACAGCT TGATGAATGG 780  
 AAACAGTTGT TAGATATCGT TAATAATTTA GATGGTAAAA TTACAATTGG TTTAGTAGGT 840  
 45 AAATATGTTA GCTTACAAGA TGCATATTTA TCAGTTGTTG AATCATTGAA ACATGCTGGA 900  
 TATCCTTTTG CCAAGATAT TGCATTAGA TGGATTGATT CAAGTGAAGT AACAGATGAA 960  
 AATGCAGCCG AATACCTTGC AGATGTCGAC GGTATTTTAG TACCAGGTGG ATTTGTTTC 1020  
 50 CGTGCAGATG AAGGTAAAA TGTGCAATT AAGTATGCTA GAGAAAAACA TGTACCATTC 1080  
 TTTGGTATTT GTTTAGGAAT GCAACTTGCA ACAGTTGAAT TTTCAAGAAA CGTATTAGGC 1140  
 55

|    |            |            |            |            |            |             |      |
|----|------------|------------|------------|------------|------------|-------------|------|
| 5  | TTACCAGAAC | AAAAAGATAT | CGAAGATTTA | GGGTGTACAT | TACGCTTAGG | CTTATAITCA  | 1260 |
|    | TGTTCAATT  | AGAAGGCAC  | ATTGGCCAA  | GATGTTTAGT | GTAAAGCGA  | AATTGAAGAA  | 1320 |
|    | AGACATCGT  | CTCGTTATGA | ATTTAAATAT | GACTATAGAG | AACAATTAGA | AGCAAATGGT  | 1380 |
|    | ATGGTGATT  | CTGGTACAAG | TCAGATGGA  | CGTTTAGTAG | AAATGGTAGA | GATTCGCAGA  | 1440 |
| 10 | AATGTTCTT  | TATTGCTGTT | CAATTCACC  | CAGAATCTT  | ATCTAGACCA | AATCGTCCG   | 1500 |
|    | ACCCGATTT  | TAAATCATT  | ATTGAAGCT  | CATTAAAAAT | TCAACAAAAT | AAATAAATTT  | 1560 |
|    | GCTAATAAAA | CCGGTACTTT | CATTGTTAAA | CATTGAAAGT | ACCGGTTTnT | CGTATAAATTT | 1620 |
|    | TAATATATG  | TTAGTGACAA | GGTATGAAAT | AACAATAGTG | ACITTTATAA | TTCTAAGTCT  | 1680 |
| 15 | CTTGTCATT  | CAATCATTG  | TGTATAAATG | TCATAGTATA | CATAATTCAA | TGCCATCGCA  | 1740 |
|    | TGTGGTYGGA | CAATCTTATC | GTAACTCTCA | GTGTAGACTA | TAGGTCCTGG | TGTAGATAAA  | 1800 |
|    | TCGATAAAAT | GTACGAGATG | ATCAGGGAAA | TCATCTGTTT | TAGGTTTGTT | GCTTATTAAG  | 1860 |
|    | ACCACATCGA | TATCTAAGTC | GATAAGTTTT | TGAATATCTA | ATGCAACTTG | ATyATTATAA  | 1920 |
| 20 | AATGGTGCGA | ATAATAATAC | ACGATCAGTT | GAGTCAATTT | CTTTAAwktT | TTTAATAGCG  | 1980 |
|    | TAAGTTTnCG | GCTAG      |            |            |            |             | 1995 |

30 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1107 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

|    |            |             |             |             |             |             |     |
|----|------------|-------------|-------------|-------------|-------------|-------------|-----|
| 40 | ATTAGTGATG | AAAGTCAGAT  | TGAAGCTTTA  | TTAACAGCTG  | AAAAATATTC  | AGAAATGATT  | 60  |
|    | GGTGAAATAT | CACCGGTGAA  | CTCCTTTAATC | TAAAGATTGAG | GAGGTTATTAT | TTAGTCTGAG  | 120 |
|    | AAAAATAAGT | ATATGAAAAA  | AAAATTATTA  | GGTACTATTA  | TTTGGAGTAT  | TGCTACATTT  | 180 |
| 45 | TATTATTCAA | GAATGATGGA  | AATAATGAAT  | TTAGCTATTT  | TAAAAATAAA  | AATTGGGGGA  | 240 |
|    | AGTTAATATG | CTAAACATTC  | AAGACGTTAT  | CATNCTTTCT  | AAAAAGGAGC  | AAAAAGCATA  | 300 |
|    | TAACCGTTTC | GTAGAATCTG  | TAGAAAAACGG | TAATTTACCA  | GTACTACCATT | GTATTGAATAT | 360 |
| 50 | GGATCTAAAA | GAGATGCAAG  | AGAAAAACAT  | GAACACAGAT  | AAGATTGGTG  | GAATGCCATT  | 420 |
|    | TTTTAAATCT | TTTTAAAGATA | TACCAATTAGA | TAGAAATTAAT | GTACCAATGG  | TATTTCTAGC  | 480 |
|    | ACAGATTAAT | TTGGATGATC  | TTCACAGACA  | ACAGAATTA   | TTTCTCTGAA  | AAAAAGGGAT  | 540 |



# EP 0 786 519 A2

AAACAATATA AACTCAAGGC TTGTTTATAT AAAAGAGCCA ATTACAGATT TATCACTCGA 660  
 5 AAATATTCAA GCGCATTGA AGTCATTAGA TGCTGATAAT GAGGATATCC CGTTTCAGTGG 720  
 AGCATTTCCT ATAGAATTTA GATTGTCGAA ACAAACATATT ACATGTACTG ATTATAAGTA 780  
 CGATGAGGAC GTGCTTGCAAT TGTGGAATAA AGTCAATCCA TCCTTCGCGC TAAAATCAAT 840  
 10 GTTTGGTGGT TATGATGAAT TGATGGAACC TGTGTGTAmC AywTTTACTG CTAAGGAACC 900  
 ATTTAATCAA CTGTGTGGTT ATCCATATTT TGACCAATAA GATCCAAGAA CGAACGATCA 960  
 AGAATCGAAA ATGTATGATA GAGTCTTACT GCAAATGAT TCTACAAGAG ATGTAATTC 1020  
 15 TTCGATTATA TGGGGTGATT TAGGTATTGc CAATATCTTA GTGaAATCTA CTGrACCTTG 1080  
 aGGcTatGaa GTTTGAATG ATTACCT 1107

(2) INFORMATION FOR SEQ ID NO: 639:

20 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 904 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 25 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 639:

30 ATTCATATTA TTATAAATTA TTCTACACC ATCCCAATTG AGTTGTTTTT CATAATTTAA 60  
 ATGTAATCC ACTAACTCCC TACCAATTTT AACAAATCCA TATACATCTT TTAATATCGG 120  
 TATTGCGGGA AAACCTTTAC TCAATCACT TGAATATTG TACACATAAT ATTTATGATG 180  
 35 CAAATTTGCA TATATATAAT ATACTATCTC TTCTGAATTA AGATTTATTT TCTTTTAA 240  
 AGAATTAGGA AATATTATCT ACArGCCTCA AACTATCTTT ACCTTTGcAT GTAGCAAGC 300  
 CTTkGCCATT ACCAATAAAc TGGAAATTAG GTAATATGTC CGTGATCATA GCCGAGAATT 360  
 40 CTTTATTCAT TCCCTGTCCT TGATATATAA TCACCTGTCC AGTATTCCTC ATTATATTAT 420  
 AATATCTACT TGGCATTTCc ATAATATTTT TGTCGTACAC TATCCATTTT TTTGTAAATG 480  
 GTCTATGCAT AAATTTAACA ATTCTCTCTG GATTAAATTGA AATATTTTTT CCTTTAGAAA 540  
 45 ATTTTTGGGT AAGTCCTCGT GTCCAACta TaAATGTTTC ATCTTTGTTT ACTAAATTTA 600  
 TACGTTCTCT TGAATCTAAG ATATCAATTA ATCTATCTAT TTCAGAATTA TAGTTATCTA 660  
 50 CAAGTAATTT TGCATTACT AATGCTTTTT CATTCGAAAA ATTTGTTACC CAATTATCTC 720  
 TTGCTGaATT TACTCCATTA AATkGAKCTA AATATATAGA ATTTTCAATA TCCTTGGAA 780  
 CATACATGG TAAATAATTC CCCATAGTTT ATGCTCTGGT GATTAAATCCA ATCATGGGG 840

55

# EP 0 786 519 A2

AAAG

904

(2) INFORMATION FOR SEQ ID NO: 640:

5

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 436 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 640:

|    |   |     |
|----|---|-----|
| 15 | CGATGTCITT ACTATTAGAC TTAGCCATTG GTTTCACCTC TCCAAAAATT GTAAATGTGT | 60  |
|    | ATCATCAATA TGAAGTTAC ATAAACTGA CATATTCTT TAAATATCA ACGCCATTGA     | 120 |
|    | TAACTTCCTG TTITAATTGA TACGCTGTAA CAAAATACTA TAGTTAGTGC TTACATGTAT | 180 |
| 20 | ATGTTAAAGC AAGCAGTGGT AAATGTAAAT TATAATTATT CATTAACTTT GCAATATATT | 240 |
|    | AAATCTTTTA TTCATAGAAG ATAAATATCA AATCAATCAT AATTATTGTA CAACAAATAG | 300 |
|    | CTAACGATTG TTTTAATCTA CATTGGCTT ATAGCATTTT AAACCTATAC TCTATTTTGA  | 360 |
| 25 | TACAATATAA GTGTAAATC AATCATAAAA AGGATATTCA ATATCTGCAT CCAAGAAAAA  | 420 |
|    | CATTACAATT ACCTTT   | 436 |

30

(2) INFORMATION FOR SEQ ID NO: 641:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 442 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 641:

|    |  |     |
|----|--|-----|
| 40 | GTTATTAAT TCAGAGTGGT AGCAATTAA AGTTATCAA GAGTTAAGAT GAATTTAATT     | 60  |
|    | CATGAACACG TCTATTATT TTATTAATTGT AGCAATAAA GCTTTACATC AAGGAGGTAA   | 120 |
|    | TTAAATATGT TCAAAAAATA TGACTCAAAA AATTCATCG TATTAATAATC TATTCTATCG  | 180 |
| 45 | CTAGGTATCA TCTATGGGG AACATTGGA ATATATCCAA AAGCAGACGC GTCAACACAA    | 240 |
|    | AATTCCTCAA GTGTACAAGA TAAACAATTA CAAAAAGTTG AAGAAGTACC AAATAATTCA  | 300 |
| 50 | GA AAAAGCTT TGGTTAAAAA ACTTTACGAT AGATACAGCA AGGATACAAT AAATGAAAAA | 360 |
|    | TCTAATAAAT CTAGGAATTG GGTTTATTCA GAGAGACCTT TAAATGAAAA CCAAGTTCGT  | 420 |
|    | ATACATTIAG AAGGAACATA CA   | 442 |

55

EP 0 786 519 A2

(i) SEQUENCE CHARACTERISTICS:

- SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 2472 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 642:

| id  | seq         | seq desc   | seq id no. | seq                               |
|-----|-------------|------------|------------|-----------------------------------|
| 0   | CCAATTCTGG  | TATGAATTAT | ACAGATAAAT | CnGCGCCCGG AGGATCATT GCTTATTTAA   |
| 10  | CTCAATTCTGG | TGTGGATAAA | TGSATGAAT  | AAGGTTATAT GGCATAAGGA GAACATTTTA  |
| 15  | ACTACTGCCA  | ATAACCGAAG | ATATATTTAT | CAAGCTGGAA CTTCATTAGC CAACCTTAA   |
| 20  | GTTCGGGAG   | CAGTACTGTT | AATCATGTAT | AAATATCATC TTGAAAAACA TCCAGATAAA  |
| 25  | GCATGTGAAT  | TGTTATATCA | GCATGGGACA | TCTAAGAATA ATAAACCATT TAGTAGATAT  |
| 30  | GGGCATGGTG  | AGCTTGATGT | GTATAAGACA | TTAAATGTAG CAAATCAAAA AGCAAGTTAA  |
| 35  | TAAATCAAG   | GAGTTTTTGA | TTATGGCAAA | ATTAGTTACT GAAAAACATT CGAAGCGGTT  |
| 40  | TAAAAATCAA  | GATGTATTAA | AGCATATTAA | TATCATTTTA GAAAAAACG AAGTTTATGG   |
| 45  | ATTACTTGGT  | ATTAATGGAG | CCGGTAAAC  | GACACTTATG AAAATTTATAT GTGGCATACT |
| 50  | TCAACAGAT   | TCAGGGGAAA | TTAAATTAGA | TAATAGACCA ATGACACGAA ATGATTTGCA  |
| 55  | CAAGTGGTG   | TGCTTATTG  | AAACACCTGC | GACATATAAT CATTTAAAGTG CACAAGATAA |
| 60  | TTTGAATAAT  | GTGTGTTTAA | ATGAAAGCGT | TGATTTCaGC GAAATTAATA GTGTTTTTAA  |
| 65  | CTTAGTCAAT  | TTAAATGTG  | ATAAAAAGAA | AAAGGTTAAG GACTTTTCTT TAGGTATGAA  |
| 70  | ACAAAGACTT  | GGAATTGCAA | TGCGaTTAAT | TAAAAAGCCA GAAATTTTAG TATTAGACGA  |
| 75  | ACCATCTAAT  | GGTTTAGACC | CATATGGAA  | CCAAGAACTT AGAGAACTTC TAAAATTATT  |
| 80  | AACAGAACAA  | GGTACTAGTA | TTATTATTTT | AAGTCACATT TTATCTGAAA TCCAAGTTTT  |
| 85  | AGCAGATCAT  | ATCGGTATTA | TTCAATGAGG | TGAGCTAAAA TATCAGCAAA GAAATAACAA  |
| 90  | AGATGAAAA   | TTAGAAGAGA | TATCTTCAAA | AAATACGAAA GGTGATTACA AATGATACAT  |
| 95  | TTAAAGATAG  | AAGGTATCAA | ATTTAAAAAA | TCTTTTCAGTA TGTATGTTTT ATTAATAAGT |
| 100 | CcGcTGGTAT  | TTCTTTGTTT | TGCTATTTTC | ACAGTCTTAT TCGCCAAAAG TAATACGGGA  |
| 105 | ATCCGGAATA  | GTGTGTCACC | ATATATACAT | TTACTATTTA ATATTTGGCC AATTGCTTTC  |
| 110 | ATCCCGATG   | TATATGTAT  | GGCTTGTAAT | TGCTTTATTA AAATGTAAAT GAGAAATAAA  |
| 115 | TCATTTAAT   | ATTACTTAAG | TAATAATTGG | TGCGATTACAA AAGAAATAAG AGCAAGAAGT |
| 120 | TTCAATTTAT  | CAATAGCATT | TTTGGTACAT | TGCTTTTTIAG TATTATTAT TGCTTATATA  |



CACAACAACC GATACAATTA CTTTTCATG ATAATAATCA GTTATACGGT TTTGTTTATC 480  
 CAATTGTAGA TAAAAAAGAA TTAAAAGATA AGTTTAATAT TAACAATAAC ATTTGGATTG 540  
 5 CTAAAGTTGG GAATGGATAT TGTATTGCCA ATTTGAAAGA AGACAAATGG ATTTATATTG 600  
 AATTGTAGGT GTAAAGATGC TAGATAATAT TATTTATAT TTAAA 646

(2) INFORMATION FOR SEQ ID NO: 644:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 426 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 644:

20 TAATATCGGA ATTGATAAT GAAGATATCT AATTTTTTAA TATTCGTAGC TTTTATTTTT 60  
 CTACTTATTT TAAGTTTATT TTTAATCTTA CAAATGACAA ACCATTAAAA GTAGCATCCC 120  
 AACATCAAAC AAAAAAACAA TTCATCAAAT AAAAATCGCT ACAAAACCAA GTCAITAAAC 180  
 25 ACGCAATAAT TAAAATTTTC CACTCATTAT AATTCGTAAT TCCAAATGTC GAATTCGGAA 240  
 AACCAACTC CAAATTCCAA AAACGCAACT CCAAAATTAA AAGCATTGCC CTACCATTCG 300  
 30 GGAATGCTT TTTACATACT GgATTACTCT GTCATTAAATG ATTTTACAAC GgGAAACCAT 360  
 GTCGTGATGT ATGACCaaAG TAGCGTCGCT AtCaTaAgGt GGTCGGATC TTTATTGGAT 420  
 AAThAT 426

(2) INFORMATION FOR SEQ ID NO: 645:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3241 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 645:

45 nTTATTATTACC CACAACATGT TGGGACATTA GGTAATGGG TACCTTATTT ACTTGGTATT 60  
 GTTATTGTTAG GTATGGGATT AACAAATTACA CCTAATGATT TCAAAATGGT CTTTAAAGCA 120  
 50 CCTAGAGCAG TAATTATTGG TGTCTGTCTA CAATTCAGTA TTATGCCCCAC ATTAGCATTT 180  
 ATAATTGCaa AGTCTTTTCA TTTACCACCT GATATTGCTG TTGGCGTAAT ATTAGTTGGA 240  
 TGTGTGTCGG GTGGGACATC AAGTAATGTA ATGAGTTATT TAGCCAAGC TAACGTAGCA 300

|    |   |      |
|----|---|------|
|    | ATATATCTAT TTGCAATGA ATGGTTGGAA GTATCTTTCG TGAGTATGTT GTGGTCAGTT    | 420  |
|    | GTTCAAGTTG TATTAATTCC AATTGCTTTA GGTATTGTTT TGCAAAATTAT TAATCGTAA   | 480  |
| 5  | ATTGCTGAAA AAGCTTCTAC AGCTTTGCCA ATTATATCAG TTGTGCTAT TTCATTAAAT    | 540  |
|    | TTAGCAATAG TTGTAGGTGG CAGTAAGCAC CAAATCTTAA CTACAGGATT ATTAATATTT   | 600  |
|    | TTAGTAGTTA TTTTACATAA CGTATTAGGG TATACGATTG GATATTGGTT AGCTCGTCTT   | 660  |
| 10 | TTAAAATTAG ATCGACAAGA TCAAAAAGCA GTCAGTATTG AAGTTGGAAT GCAGAACTCT   | 720  |
|    | GGTTTAGCTG TGCTATTAGC aGCAATGCAT TTTAATCCAA TTGCAGCAGT ACCAGGCGCA   | 780  |
|    | GTGTTTAGTT TCATTCATAA TATAACAGGG CCTATTTTAG CAAAGTATTG GTCAAAAAAG   | 840  |
| 15 | TTATAATTGC ACTAATAGAA TGAAGTGGTC ATCGGACTAT GTTAAGCTTT GATAAAGAGA   | 900  |
|    | AAAAATAGAG GAGTAAATAT ATGTATAGAG CAGTTATATT TGATTTCGAT GGAACAATAA   | 960  |
| 20 | TAGATACGGA ACAACATTTA TTTAATGTTA TTAATAAACA TTTAGAGATG CATAATGCCG   | 1020 |
|    | ATCCTATAAG CATTGATTTT TATCGTTCTT CTATTGGAGG AGCAGCTACA GATTTCGATG   | 1080 |
|    | ACCATTTAAT TAAAGCGATT GGTTCGGAAA ATAAAGATAA ACTTTATGAA GAACATCATC   | 1140 |
| 25 | TTACTAGTAC AACATTACCG ATGATTGATA CGATTAAATC ATTGATGGCA TTTTAAAGC    | 1200 |
|    | AACGTCACAT TCCTATGGCA ATTGCCACAA GTAGTGTGAA AGCGGAAATA ATGCCACCT    | 1260 |
|    | TTAAAGCATT AGGCTCTAGC GATTATATAG AGGTAGTTGT TGGTAGAGAm GATGTTGAAC   | 1320 |
| 30 | AAGTTAAACC TGACCTGAA TTATATTTAT CTGCAGTACA ACAATTAAT TATATGCCGA     | 1380 |
|    | CACAATGTTT GGCTATTGAA GATTCTGTAA ATGGTGCAAC AGCCGCGATT GCAGCTGGAT   | 1440 |
| 35 | TAGATGTAT TGTTAATACG AATAAAATGA CAAGCGCACA GSACTTTTCT AATGTAGATT    | 1500 |
|    | ATGTAGCAAA AGATATTGAT TACGATCAAA TTGTAGCGCG TTTCTTTACG AAATAGGAGG   | 1560 |
|    | CGTATCATGA TGGGTTACAT TATATTGTTT TTTCTAGCTG GTCCAGTAAT TTTAGGCGTT   | 1620 |
| 40 | GGAAATTGGT TGATTGGTCC TATATTTAAC AAACAGACAC CATTTCGCGT GCAAGTAAGA   | 1680 |
|    | TCITTTGTTG KTGGKTCaAT GTTTTACTTA ATACTCGCAA CAATTGGCTA TTTTTTACTA   | 1740 |
|    | TTACAAGGTA AACTTTAAGC AGAAAACCA CTTACCTCAT TAAATGGAGC ACCATATGTA    | 1800 |
| 45 | TGTGAAATGG TAGAACGTTT ATGTTTATGT ATGAGATAGG GTGGTTTAAA TAGTTACATA   | 1860 |
|    | TATTTTAATA ATAAGCTCAC GATGATAAGT ACAATTAAGA TAATATCTAT GCCTACCATA   | 1920 |
|    | ATTGTAGTCT TTGTGCAIT ACTTCCTTGT TCTTTTGCTG ATTTTCATAGC ACGGTAGTTT   | 1980 |
| 50 | GGCACAAAGC TAATAATTAG TAAGATTAAAT ACAATTACAC CAATTAAATGC TGTGTGCATG | 2040 |
|    | ATGAACGACC TCCTTTATTT TTTTCAATCA ATTCCCAAAT AAACGTAGCA ATCACACCGA   | 2100 |

|    |  |      |
|----|--|------|
|    | CAATAATTAA TGCAATCGGT AAAGTCGTAC CGAGTTTAAAT CTTCGCTCT GGAGAAATTAA | 2220 |
|    | TAATAGTAAA TACTGTGAAGA CAAATGAGTA TGAAAGCAAG TGTTGCAATA ATAGTCTTTC | 2280 |
| 5  | CAACTAAATA TAGGATGTCA GGTTTTTCCTA TACCGATATA ATTTATGATG AAAAATGCTA | 2340 |
|    | CAGCAAAGAG TACCGATATT TTTGTAGCAC GTAGCAGTAT TTGTTTAAAC ATTGATATAC  | 2400 |
|    | TCCTTTTIAA TATTATTAAA ATTATATCAT AATTACCAAG AATAGCTGAA GTTGATGTG   | 2460 |
| 10 | ACTCAACGGT ACTTGAGCAA CTTTTTAAAT TTTTIAAGAA AATCACAAAA TAAATTGTTG  | 2520 |
|    | CAAAGTGC AAGCCTGCT ATAGTAGTTC TGTAACGAT TGCATCGTAT GCAAAATATA      | 2580 |
|    | ATGTACCAAA ATCGATAATT TATAGTATAA TTACGGCAAT AAGTTTTTTT ATGGATTAT   | 2640 |
| 15 | TTAGTATCAA TCAGAGATGG GGTAGAAGT TATGGAGAAC AATGAACAC AAGGGGGATT    | 2700 |
|    | GAGTGCCCGT CAAATTCAAA TGATTGCACT TGGTGGTACG ATTGGCGTGG GCGTTTTCAT  | 2760 |
|    | GGGTGCGACA AGTACAATTA AATGGACAGG CCCATCAGTT ATCCTTGCAT ATTTAATTGC  | 2820 |
| 20 | GGGTATCTTT TTTTITTTAA TCATGAGAGC AATGGGGGAA ATGATTATT TAAACCCCTAC  | 2880 |
|    | AACAGGATCA TTTGCAACAT TTGCAAGTGA TTATATACAT CCTGCACGAG GTTATATGAC  | 2940 |
| 25 | AGCATGGAGT AATATATTCC AATGGATTGT AGTTGGTATG AGTGAGGTCA TCGCAGTAGG  | 3000 |
|    | AGAATATATG AAGTTTTGGT TCCCGAATT GCCAACTGG ATTCCTGGTG TTATTGCTAT    | 3060 |
|    | TTTATTATTA ATGGCAGCGA ATTTATTCTC GGTAAAAGCG TTTGGAGAA TTTGAATTTTG  | 3120 |
| 30 | GTTTGCTTTA ATTAAGTTG TAACAATTAT TTTAATGATT ATTGCTGGTT TTGCTCTTAT   | 3180 |
|    | nnTCTnTGGT TTTGAAATG GTGCCATGC GGTAGGTATT TCTAATCTAT GGACAAATGG    | 3240 |
|    | C  | 3241 |

## (2) INFORMATION FOR SEQ ID NO: 646:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1311 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 646:

|   |     |
|---|-----|
| AGGCGTCAAC TCAGATGGTT TAATAATTGC CGTATTACCT GCTGCAATAG CACCGATTAA | 60  |
| AGGT:CGAaC ACTAGTtGAA AAGGATAGTT AAATGGTGCA ATGATCAAAA CTGTTCCATA | 120 |
| AGGTTCTTTT TTGATATAGC TTTTGTGG AAATAAATAT AAAGGTGGT CTACATKTT     | 180 |
| TGTTTIAAGT CAGTTTTTAA GTTCCTTACG GGCAATTTTG ATACTTtTCA AAGTTATGCC | 240 |

|    |  |      |
|----|--|------|
|    | AATATCGCTC TCGTATGATT TAATAGCITT GCTTAACTTC TTTAATTGCT CTITTTCTAAA | 360  |
|    | ACTAATATCT TTAGTTTGGT GTGTATTGAA AAAAGCTTTA CTGTCATAAA ATTTTTGCTC  | 420  |
| 5  | AATGATATTC ATAATGAAA GAACCTCCTT ATATGATTAT TTTGGAAAAA GCGATTAAAT   | 480  |
|    | GATTTGAATG TTGTGGCCGT TAATTTTAAA TGGTCTTTTCG AATTATATAT GTTGAAGATT | 540  |
|    | GAAAATAGAG CGATGAATCG TGTACATAAT AATATTTATA ACTTTAATCA TAACGAAAAA  | 600  |
| 10 | GGTAGGAAGA AAACAAAAAT TTATACTCAA CATCGCAAAT ATTTTAAGAA AATGTAAGA   | 660  |
|    | CAAAAGGGGA ATTGTATAGA AATCACTAAT CTGTGGGTTA GGGTAGCTAA AGGAATAAAA  | 720  |
|    | ACTACTATTG AAAAAGGGTT GTAAATTAGT CAAACGTAAA TAAAAACAG TTCATTGAAA   | 780  |
| 15 | GTGAAATAAA TTCTACTTTA ATGAACGTGT AGTTAAATAC AACATGTCTA TAATTAGACA  | 840  |
|    | GTAATATAGT ATTATTTTGT TAATGCTTCA GTGATTTGAG GTACGATTTG TTTTTTTCGA  | 900  |
| 20 | GAAACGACAC CAGATAGAAA GGCCATGTCA TCTTCTAATT GAACATTGAA TGTTCGCCAA  | 960  |
|    | CTTTATCTTT TTGAGCACCT ACAACTAAAA TTTTAGAATC ACTATTAATG ATGTCAAGTA  | 1020 |
|    | CAACAAGTAC AAATAAGTCA TATTTTCTT GTGCATTAC AGCTAACATT TCTTTTCTA     | 1080 |
| 25 | TATCTTCTTT ACGATTTAAC ACTTCGTCAA GGTCAACAGC ATTAACTTGT GCAATACGAG  | 1140 |
|    | TCACATAGTC ACCCATAGTA AATGATTAGT CATCCATGTT TAATAAGAAT TCAACTGATT  | 1200 |
|    | TATCAGTTGT TGAAGCACCT GCITTTAACA TATCTAAGCC GTACTTTTGA ATATCAACTT  | 1260 |
| 30 | TAGCAATATC TTTAAATICT TCAGCTGCTT TAACATCTTG TTGTGTACAT G           | 1311 |

(2) INFORMATION FOR SEQ ID NO: 647:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1498 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 647:

|    |  |     |
|----|--|-----|
|    | CATATATCTT TTCATTCTTT TACGAGATAC TTTACCAGAG GATTTAGACT TCATACGCTG  | 60  |
| 45 | ATCCATATGT GCTTGGGTTT CAGaATGTCC ACAAACACAA CGATATACCG CTTCCTTCCC  | 120 |
|    | TTTACCAAAC AACGTTAAIT TCTTTTACA GTTTGGACAT CTGCAATTTG TTTTGCCTG    | 180 |
|    | TACATTCTTT TTGCTCTTAC AAGATGGATC TTGGCACACA AGCATCTGAC CATTTTGTAGT | 240 |
| 50 | TTTAACTTAA ATCATGAATT TACCACAGT TGGGCATTCT GTGGTGTGTA AATTATCGTG   | 300 |
|    | TTTATATTAA CGATCACTAT TTTTAAATCCC ATTTACAACA TCTTTCGTAA AATCTTTCAT | 360 |



## EP 0 786 519 A2

|    |  |      |
|----|--|------|
|    | CCATTGTGcA GTTAAAGAGT GCGACGTTAA TTCTCTGGT GCTAATTCTA ATATTGTTT    | 480  |
|    | ACCTTTTGAC GTTACTTTAA TTTTACCGTC TCTTGATTCA ATGGCATTCA TATTAAATAA  | 540  |
| 5  | TTTATCGATA ATGTGCGGCC TTGTGCAAC TGTGCGGATA CCACCTGTTT GTTTTAAAGT   | 600  |
|    | TTGCGCATAT TTTTATCTCT TCAATTGAAT AAGGTTCTGA GGGTTCTCCA TCGCTTTTAA  | 660  |
| 10 | TAAACGAACCT TCATTAAAAA ATTCTGGAGG TGTGTGTTCA TGTTCCTTAA TATTGTGTTT | 720  |
|    | TGAAATCTTC ACTTCATCGC CTTCTGAAAA AGGCTGTGTC ATCTCTGTAA TAGATTCAAC  | 780  |
|    | TGTGCTAATA GATTTAAAAA CTA AAAACAGT TGTTACATTC TCTTCAAAA CAAATGTGTG | 840  |
| 15 | CCCTGCAACC TCTAAGGTIA CAGTTATCGC GTCATACTCG TCGGAGGCA TTAAGCTTTC   | 900  |
|    | TAAAAAACGC TCGACAATCA TATCGTATAA CTTTAATTCT CTATTACTTA AGTCTGACAT  | 960  |
|    | GACAGGCTCT ACTTCTGTAG GAATAATTGC ATGGTGATCA GATACTTTTT GATTATTAAA  | 1020 |
| 20 | TATCGACATT TTTGATGAAA ATGTTTTAGA CATTAATGGG CGTGCTTGGT CTTTATATGT  | 1080 |
|    | TGTTGCCATC GTCACCTGAA TACGTTCTTT CATAGTATCT ACCATATCAG TTGTAAATA   | 1140 |
|    | GTTTGAATCT GTTCTGGAT AGGTTACGAC TTTATGCTC TCATATAAGC TTTGAAGTGT    | 1200 |
| 25 | ATTCAATGTT TCTTTAGGTC CAATTTTATA ACGTCTATAC ATATCTTGTT GTAAATCTGT  | 1260 |
|    | TAAATGTGAC AGTGATTGCG GATACGACTT CTTATGTTTA GTAGCAACAG ATTTAATCTT  | 1320 |
| 30 | ACCATCGACA TTTTTCAAAT TATTAACCAT CTGTTCTAAA GTTTCCTTAA TGGCATATCG  | 1380 |
|    | CTGATTTGaw TCTAGCTGAA AATCAAACCC TTTTACGGTT AATGATAATG TAAAGTATTG  | 1440 |
|    | TTGTGnTTG AACTGATTAA TCTCTTGTG TCGTGTAAT TTACTAAATT GAAACGgn       | 1498 |
| 35 | (2) INFORMATION FOR SEQ ID NO: 648:                                |      |

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1044 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 648:

|    |   |     |
|----|---|-----|
|    | TAAACAATT CGTTTTATTA ATACAGTTTG TAAAGTATT CTCTAGTAC TAATAGCTAA    | 60  |
|    | TGTAATGATA GAAATACTA GTGTTATTAA AGATTGTAAT AAAATAAAG AAAGTGAAG    | 120 |
| 50 | ATATTGGAAT GTATTAGATG ATTATTACAC GATTGAATTT GCACCTTATC ACGAAACAAA | 180 |
|    | ACAAAGTTTG ATTGATAATA TGGTGCATC AGAACAAATTA GTAAGGCTA GTGAAGCAGA  | 240 |
| 55 | AAATAATGCG ATTTTATTCA AACCAAAGGG TGACTCGGT GACAATGACA ACTTTTCGCC  | 300 |

TCAACCTGAT ATCCGATAA AAAATCAAAA AAATAATGTC GAAGTAATTA TTCCACAAAA 420  
 GTTTCATGCA ATGCGTAATG AAATCAATCA AGCATATCAT TCATGGTTTG AATTGTGACA 480  
 5 AAATAAAAAA AATAAGAGA ATAAGTTATC TATACAGTTT ATCAACAAAA ATGATTGTCG 540  
 AATTTTTTCA TTGATGTCAC GAGATAGTCG CCATTGTGCA TTTATAGAGG CGCCAATCAT 600  
 TGTGAATGTT CAGGCATCAG ATTTATCGAA TGATTTTtAT TATGCCATGa TcAGTCaAGG 660  
 10 CGGGTATTtA TTCaAAAAAT ATGaCGCGCT AGTAAAAAAT ATTGGAAAAAG TATCCATCTT 720  
 GATGGGGAAA TCCAGTGGA TAACCAATTA TAAAGATAGC GTGATGGAAA TGATATCATGA 780  
 AAACAATTG AAATTAACAG TACTCAACTT TTCACAAATC ATTATCGAAA TCATTTTAAAT 840  
 15 AATTATTATT TTATTTGATG TGAAATATTA TTTTGAACAG CATCGAAAAAT TACTCGTAAT 900  
 CAAAAAGCTA TATGGTTATT CAACATTAAG AGCCAATTAC CAATACTTAT TAATAAATAA 960  
 20 TATAGTTGTT ATTTTTATTG GAATATTGAC GAATGTAAAT TTACATTCTC ACTATATAAT 1020  
 GATGTTATTT GCAACGATTC TTGT 1044

(2) INFORMATION FOR SEQ ID NO: 649:

25 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 399 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 30 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 649:

35 GAACATATTG GGTATGCAA GGrGGTCACT CTTCAACTT ATAAACAACA TTTTAATAAT 60  
 GTAAAGTTTA ACCAGCTAAC ACTTTTGTTA GCTGGTTTTT ATTTTCCTTC AATTTTAAAA 120  
 TGGTTAAGTC CCTTCTATA TCTTATAAGA CAATCAATTAT AATCAATTCA ATTAATACAT 180  
 40 TAACAACCAC AACTAATAAA TATAGTAAC TCAAAATCCA TATTTATGTC TAAAGATAAT 240  
 CTCAATGTTG TTCACGTCAA TAAATTATC CCTAGTITT TAAAAATTGT ACATGTTTAA 300  
 ACAATCAAAA GTGTACATTA TTAAATTATC ATTTCCAGTT AGATTTAGAA AACATTCACA 360  
 45 CCACGCATGG ACCAACGTAT TCGTCTTcAT TcATTTTAG 399

(2) INFORMATION FOR SEQ ID NO: 650:

50 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 747 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

55

# EP 0 786 519 A2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 650:

|    |  |     |
|----|--|-----|
|    | GTACTAGGTT CTAGATTCC CTACTGTAGT CATAGATATT TTCCAAACAG AGCCTGACAA   | 60  |
| 5  | CTTTTCATT TTAATAGTGA TCACATATATT AATCGCTATT GCAATTTATG CATGTCGATT  | 120 |
|    | TGTTTGGGTT TATTTCTGGT ACAAAGATT TTATTTCCCG AAAAATATAC AATCTTATCT   | 180 |
|    | AGACGAGGAA CATGATTAC ATGAACACCC ACCTTCTCGA GTGCGTTACG CATTATTATT   | 240 |
| 10 | GACCATGTGT GGTATTCACG GTACAATTTC ACTTTCAATG GCACCTTACAT TACCATTTAT | 300 |
|    | CATTACAAAA GGACAAGCAT TCGAATACCG TAATGATTTA TTGTTTATTG CATCTTTTCT  | 360 |
|    | GGTATTAATT AGTTTAATCT TAGCGCAAAAT TGTTTACCT TTAATTACAC CATCTGCCGA  | 420 |
| 15 | AGATACTACT TTTAAAGGTA TGACTTATCA ATCTGCCAAA ATTTTCATTG TTCAAAAAGT  | 480 |
|    | GATCCAGCAT TTTAAAAACG AAAGTAAAAA AGACAAAAAC GATACAAATT ATCGCCCACT  | 540 |
| 20 | ATTAACCAAA TACTATGGAG AATTGTTATT TTTATTAAT TCAGAACCTG ATAATCAAAA   | 600 |
|    | TACAAAGAA CTCAAACGTT TAGAAGATAT TGCaaaaGTA ATCGAAACAT CTACACTTGA   | 660 |
|    | CGGTTTAATT GATAAAGGTA AGGCAACATA TCAGGATATT AATAATTACC GCAATATTGT  | 720 |
| 25 | CGAATTAACA GAGACACACC GTACTGC                                      | 747 |

(2) INFORMATION FOR SEQ ID NO: 651:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1373 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 651:

|    |   |     |
|----|---|-----|
|    | ATACAATACT CTTTATTAT TCAATAAGCC ACTTCCTATA GCAAAATGTTT AACTTTAAAT   | 60  |
| 40 | ATTTTTCGAT GCTAACAAAA AATCACACTA TCATCTTTTA AAATGAAAGT GTGATTACAA   | 120 |
|    | GCAAAATCTGT AAAATTATATA AAGCAGAAAC AATTCAACTT TATCATTATG ACATTTCAAT | 180 |
|    | TAAACCTTCT ACATTATAGT TCCAAGCATC TTACACATGA ATGCAAGTAT TTAACGATTT   | 240 |
| 45 | AATTGIGACA TAGCCTGTTG ATATTGTGTT TCATTGATAT AATTTTGTG CTTCATTTTT    | 300 |
|    | TCTAAGTTCG TGCTTACAG TTGCGTAAAA TTCTCTGACA TATTATTGAT ATTATATACG    | 360 |
|    | CTAGGTGCAT TGACTTTACT AGCTAAAAA GCGCTTTGTA AAAGTGTAT GTGAGACATT     | 420 |
| 50 | GTTGTACTAT TTTTATTCAC GGTGTGTCCA AAGTAATGAT TTGCTGCGCC CTCAGCGTA    | 480 |
|    | TATTGATTAT CCCCAGGTA AATATTATTT AAATAAAAGC TTAATAATTC GTTCTTATTA    | 540 |

TCATTATCAT AAAAATAATT TTTGACAACT TGTGTGTAA TGGTACTACC ACCTTGACACA 660  
 TCTCTGTCGC TAATCGTTGA AAATAAAGCT CTAGTTGTAC CTTTCAAATC GAATCCATGA 720  
 5 TGATTGTAGA ATCGTTCATC TTCCATTGAA ATAAAGGCAC CTTTAACATA CTCTGGCATG 780  
 TTATCAGCTG ACACAAAAC TCTTTTATT TCAATTTTC TTAGTTCATC CACATTATCG 840  
 10 CGGTAGATA AAAAATACAT GATACCAATA AACAAATGCA TAATGATTAG AATGGTTAAT 900  
 AATATTTTGA ATAGTATTCG TTTACTTTTT TTCTTTTTGC GCGGTTTGCC AACTGGTTGA 960  
 TAATACGTAT TATAGTGAGG TTCGTGTTTC ATATGCTCAA AATGTTCAAT TGAGTTTGAG 1020  
 15 TACCTATCGC TCTTTTCAT GCGTTTGCTC CTTCTTTTAA AACTCACTTA GATATACCT 1080  
 TGA GTTTACC AGTACTATCA CAAATAGGCT ACACITTTTT GGAAATCAG TCCAAGGGCT 1140  
 TACAATCGTA TACGCCATCA TACTTACTTT TTTGTTTTT GAAAAATTA TAGATAAATC 1200  
 20 ATTGCAATTT TAAATATTAA TCATGTCAA TATTGTTATA TTTTATAAAA ATAAAGACC 1260  
 ATCCCTAATA AATGCCAATA GAGACGACCT TTTATTTGTT ATTCAATTTA TAAACTAAA 1320  
 ATCCATATTT CATTTCAAAC GAAATATAT AAATTTTAAC AATCGTTAAC CAC 1373

(2) INFORMATION FOR SEQ ID NO: 652:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 859 base pairs  
 (B) TYPE: nucleic acid  
 30 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 652:

GGCAGATAAT TTAGTCATTG TTGAATCGCC TGCAAAAGCA AnAACCATTG AAAAGTATTT 60  
 AGGTAAGAAA TATAAAGTTA TAGCTTCAAT GGGACACGTC AGAGACTTAC CAAGAAGTCA 120  
 40 AATGGGTGTC GACACTGAAG ATAATTACGA ACCAAATAT ATAACAATAC GCGGAAAAGG 180  
 TCCTGTTGTA AAAGAAITGA AAAACATGC AAAAAAGCG AAAAAAGTCT TTCTCGCAAG 240  
 TGACCCCGAC CGTGAAGGTG AAGCAATTGC TTGGCATTTA TCAAAAATTT TAGAGCTTGA 300  
 45 AGATTCTAAA GAAAATCGCG TTGTTTCAA CGAAATAACT AAAGACGCTG TTAAGAAAG 360  
 TTTTAAAAAT CCTAGAGAAA TTGAAATGAA CTTAGTCGAT GCACAACAAG CGCGTCGAAT 420  
 50 ATTAGATAGA TTAGTTGGCT ATAACATCTC GCCAGTCTT TGGAAAAAAG TAAAAAAGG 480  
 GTGTGACGCG GGTGAGGTTT AATCTGTTAG CmTTCGTTTA GTCATTGACC GTGAAAATGA 540  
 nATTGGAAC TTTAAACCAG AnGAATATTG GACTATTGAA GGAGAATTTA GATACAAAAA 600



TTGTGCTTCT TTTTtagcTT CTtGAACCTC TTGTGCCTCT TGTGATGTAT CACTyAAATT 120  
 ATTtGCACTT GCTTCTtCIT TTATCGCTGC TTGTtGTGCT TTCAATGCCA CTGCTTTTGt 180  
 5 TTCTTyATIT GATACAGCCA CACTTTTATC CGCTTCTGCT TGTGCTTCTC TTTTtagCTTC 240  
 TTGAATCTCT TGTGCTTCTT GTGATGTATC ACTTAAATTA TTtGCACTTG CTtCTTCTTT 300  
 10 TATCGCTGCT TGTtGTGCTT TTAATGCCGC TTGCTCATIT TTAGATTtGT TTA AAAATCC 360  
 TTCAACAGT TCTTTGTAT AGGCAACCGT TTCTTCAAGT TGC GTTTTTTCT TTTCTTCAAA 420  
 CTTTTGCGAC AGTCTTGTt CTtTGACTTT tAAATCATCT GCTTTTTGAT AAACtTTATt 480  
 15 TTA AAAATACC AACCTAAAGC C 501

(2) INFORMATION FOR SEQ ID NO: 655:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 830 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 655:

CACCTTGtCA TAATTAATTT TTTGATTTTT CATTTTACTG ATAATAGGTT CAGCATTAAAT 60  
 CATGATTtAA CCTCCACAT TTAATCATTa ACTTCTATTA TATATGATTC ATATTAAATG 120  
 30 TCAGTCAAAA AAGTTAGAAA TTCATTtTAA TGCAATTATA TTTGAAAAAT CCCTTATGTA 180  
 TCTAAAAGCA TTTTTTAAAC TTGAATTtTA AAACACTAAA CAACACATAC GTCTCTGTGT 240  
 35 CATTTTCATT TTTTGTATGT CATATATATG TTTACTTCAT TTAAATCAAT TTCATCTTAT 300  
 AATTATCGT GTATTTTACA AAAGATTGAC TTCAATTtCAT CGTAAAAAGTT ATACTTTTGC 360  
 CATTTTTTAA TGTAACATGG TGTTAGTAAT AAAAATAATA CATTGAGGTG TTTTACATGA 420  
 40 CAGCATTATT CCCTTATATC GCTTTTGAAA ATTCAAAAGA AGCCCTTGCA TATTAGGAAG 480  
 AAGTATTtGG TGCAACTGAC GTTAAACGTT TAGAAGTTGG CGAAGAACAA GCGTCACATT 540  
 TGTGTATGAC TAAGGAAGAA GCGCAAGAG CAACTATGCA TGCTGAATTT GAAGTGCTTG 600  
 45 GCGTAAAAGT GTTATGTTCT GATTCTTTTG GTCGCGCTGA CAAAATTtAT AATGGCATAT 660  
 CATTTATTAAT TGATTATGAT GTTAAACAATA AGGAAGATGC TGATAAAGTT GAAGCATTCT 720  
 50 ATGAGCAAAAT TAAAGATCAT TCTTCAATTG AAATAGAATT ACCGTTTGCT GACCAATTCT 780  
 GGGGTGGCAA AATGGGCGCT TTTACCGATA AATACGGTGT TCGTTGGATG 830

(2) INFORMATION FOR SEQ ID NO: 656: